# BIOTIN SYNTHESIS IN *BACILLUS SUBTILIS*

BY

MIGLENA MANANDHAR

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Microbiology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor John E. Cronan, Chair
Professor Jeffrey F. Gardner
Professor James A. Imlay
Associate Professor Cari Vanderpool

## ABSTRACT

Biotin is a cofactor required by all three domains of life. It is an enzyme cofactor that acts as a "swinging arm" to transfer carboxyl groups in important metabolic pathways involving carboxylation, decarboxylation and transcarboxylation reactions. Despite its importance, the biotin biosynthetic pathway has only been fully characterized in *Escherichia coli.* Our current understanding of other possible biotin synthetic pathways remains incomplete since various biotin synthesizing organisms have diverse *bio* genes that are not homologous to those of *E. coli*. The diversity in pathways lies in the first stage of synthesis of a pimelate thioester moiety. *Bacillus subtilis* represents a group of microorganisms that follow a different pathway for biotin synthesis. Genetic and biochemical studies identified *bioW* and *bioI* as two genes required for pimeloyl-CoA and pimeloyl-ACP synthesis, respectively. The question of the significance of each gene is striking due to the redundancy of pimelate thioester generation. BioW, a structurally unique enzyme, remained to be fully characterized for its importance and enzyme activity. BioW enables *B. subtilis* to use free pimelic acid as a precursor for biotin synthesis unlike the case in *E. coli.* However, the source of pimelic acid is unknown in bacterial metabolism.

My results show *bioW* is essential for biotin synthesis whereas *bioI* is dispensable. I uncovered a unique function of BioW as a proofreading enzyme of noncognate acyl-adenylate substrates other than pimeloyl-adenylate to ensure proper initiation of biotin synthesis. I also report pimeloyl-CoA as the preferred substrate of *B. subtilis* BioF unlike *E. coli* BioF. My $^{13}$C-NMR studies of labeled biotin elucidated the presence of pimelic acid in the cells and provided direct evidence for generation of pimelate through fatty acid synthesis. Decreased biotin production in presence of fatty acid inhibitors further verified my findings to establish pimelic acid as a link between fatty acid synthesis and biotin synthesis. Hence, in this Thesis, I report my observations that answer long-standing questions about *bioW-bioI* gene redundancy and pimelate source, to gain further understanding of biotin synthesis in *B. subtilis.*

## ACKNOWLEDGEMENT

I want to thank my Ph. D. advisor Dr. John E Cronan for his guidance and support throughout graduate school. Despite having numerous responsibilities other than being the Head of the Department, he always had time for his graduate students to discuss the tiniest problems we faced and to give us insightful suggestions to solve them. I cannot thank him enough for giving me the opportunity to be one of his graduate students and study the topic of my choice in his laboratory. I want to thank Dr. James A. Imlay, Dr. Jeffrey F. Gardner, Dr. Cari Vanderpool, Dr. Andrei Kuzminov and Dr. Charles Miller for their invaluable suggestions for my research. Thank you to Debra K. LeBaugh and Diane L. Tsevelekos in the Microbiology Department Office for being ever present to assist us with the paperwork and delightful email notifications about available free food. I would also like to acknowledge the financial support provided by the James R. Beck Graduate Fellowship in Spring 2014.

I want to thank my colleagues in the Cronan lab and other friends for being a positive influence on me. My special thanks to Dr. Steven Lin for his constant and brilliant advices that helped me shape my research project. Being in a big lab had its challenges but it also enhanced the sense of belonging and camaraderie among us that will stay with me for the rest of my life.

Dr. Caryn Evilia, my undergraduate advisor has been an inspiration to me since day one. I gravitated towards research and graduate school because of her. She encouraged me to pursue higher education and gave me enormous amount of support throughout my undergraduate years as my mentor. I will always be grateful for her inspiration and belief in me that encouraged me to go beyond a bachelor's degree.

I want to dedicate this work to my parents and my brother who, despite living across the globe, have always believed in me. Finally, I could not have gone through graduate school with as much ease without the loving care and enduring support of my husband, Abhishek Jaiswal. I am happy that we were both accepted for graduate school in the University of Illinois at Urbana-Champaign and were able to start our journey together.

## TABLE OF CONTENTS

Chapter 1: Introduction ..... 1
Discovery and Function of Biotin ..... 1
Biotin Synthesis ..... 2
Specific Aims of this Thesis ..... 11
Figures ..... 13

Chapter 2: Proofreading of Noncognate Acyl Adenylates by an Acyl-Coenzyme A Ligase ..... 19
Introduction ..... 19
Materials and Methods ..... 20
Results and Discussion ..... 23
Figures ..... 27

Chapter 3: Pimelic Acid, the First Precursor of the *Bacillus subtilis* Biotin Synthesis Pathway, Exists as the Free Acid and is Assembled by Fatty Acid Synthesis ..... 34
Introduction ..... 34
Materials and Methods ..... 36
Results ..... 41
Discussion ..... 45
Tables and Figures ..... 49

Chapter 4: Conclusions ..... 59
Summary of Findings ..... 59
Future Directions ..... 64
Figures ..... 67

Chapter 5: References ..... 71

# CHAPTER 1
# INTRODUCTION

## DISCOVERY AND FUNCTION OF BIOTIN

A cofactor is a chemical compound or a metal ion essential for the function of an enzyme. Biotin is a covalently bound cofactor required for the function of key metabolic enzymes that carry out carboxylation, decarboxylation and transcarboxylation reactions in fatty acid synthesis, amino acid metabolism, gluconeogenesis and other processes (67, 121). It was first discovered in 1901 as an unknown substance "Bios" that was required for growth of yeast (129). Over the following years, several researchers independently discovered biotin in different aspects. It was identified as coenzyme R, a respiratory factor essential for respiration of rhizobia (3). Kögl and coworkers associated the yeast growth factor to "biotin" that can be extracted from egg-yolk (68). It was extensively characterized as a curative factor (vitamin H) which could cure egg-white injury, a dermatitic disease, in rats (9, 49, 48) and chicks (28). Further study revealed that eating raw egg-whites facilitates the sequestration of dietary biotin by avidin, rendering it unavailable for its physiological function thereby inducing vitamin deficiency (47).

With the demonstration of the identity of vitamin H and coenzyme R with biotin (46), the work on the elucidation of the structure of biotin escalated in the 1940s. Melville, Du Vigneaud and co-workers were instrumental in deciphering the correct structure of biotin (25, 26, 87). It consists of fused heterocyclic rings comprised of a ureido ring and a tetrahydrothiophene ring that is attached to a valeric acid side chain (Figure 1-1A). The ureido ring with carbonyl oxygen is planar while the tetrahydrothiophene ring is envelope-shaped with the sulfur protruding out of the plane (22). After the structure elucidation, the first chemical synthesis of biotin was described by Harris and coworkers in 1944 (50) and a process for commercial production of biotin by chemical synthesis was patented shortly thereafter (41).

Biotin is a molecule that is required in small amounts compared to other vitamins in carbohydrate, fat and protein metabolism. Hence, unlike other vitamins, it was thought to be involved in a cellular process specific to synthesis of essential metabolites (71). Consequently, the physiological role of biotin was identified in utilization of bicarbonate for carboxylation reactions in *Lactobacillus arabinosum* (51). Lynen and coworkers were instrumental in finally linking biotin to enzymes that conduct carboxylation reactions necessary for central metabolic processes (80). In a physiologically active state, the carboxyl group in the side chain is attached to a ε-amino group of conserved lysine residue of biotin dependent enzyme subunit that enables it to act as a "swinging arm" to perform enzyme reactions (Figure 1-1B). For carboxylations, biotin protein ligase covalently attaches biotin onto biotin carboxyl carrier protein (BCCP) subunit of biotin-dependent enzymes such as acetyl-CoA carboxylase, propionyl-CoA carboxylase, pyruvate carboxylase and 3-methylcrotonyl-CoA carboxylase (121). Biotin carboxylase (BC) subunit carboxylates N1 atom of the ureido ring using bicarbonate as the $CO_2$ donor. Carboxyl transferase (CT) subunit then transfers the carboxyl group from carboxybiotin to the cognate substrate. Biotin is used as a cofactor by oxaloacetate decarboxylase to couple the decarboxylation of oxaloacetate to sodium translocation across the cell wall (72). Biotin dependent transcarboxylases from *Propionobacterium shermanii* can transfer carboxyl groups reversibly from 2-methylmalonyl-CoA to pyruvate to produce propionyl-CoA and oxaloacetate (94, 120). Hence, biotin plays a significant role in primary metabolism important for the survival of a wide range of living organisms.

## BIOTIN SYNTHESIS

The biotin biosynthetic pathway is widespread among bacteria, archaea, fungi and plants (110). Mammals cannot synthesize biotin and hence require biotin to be ingested. Biotin synthesis pathway was first studied in *Escherichia coli* by Campbell and coworkers who identified ten *bio* mutants that require biotin from growth in minimal media and eluded to at least four genes mediating the process of synthesis (20). A year later, Eisenberg and coworkers identified a four step synthetic pathway, pimeloyl-CoA to

biotin, by performing cross-feeding assays, analyzing excretion products and growth responses to biotin vitamers (intermediates in biotin synthesis pathway) of sixty *E. coli bio* mutants (111). Deletion and complementation analysis led to the first identification of *bio* genes, namely *bioA, bioB, bioF, bioC* and *bioD,* in a single operon in *E. coli* (17). This provided a strong groundwork for further extensive studies on the genetic and biochemical aspects of microbial biotin synthesis pathway. The biotin synthesis pathway has been studied in several microbes; however, the only fully characterized pathway is that of *E. coli* (75). The general pathway can be separated into two stages. Four well-conserved genes — *bioF, bioA, bioD* and *bioB* — are involved in the second stage in which the fused rings are assembled (Figure 1-2). The genes involved in the first stage of the pathway that generate pimelate thioester intermediates are quite diverse among different organisms, which will be discussed later in this chapter.

The Conserved Second Stage of Synthesis

The conserved genes involved in the second stage of the pathway were first characterized in *E. coli.* Since then, numerous other studies have shown the conservation of genes *bioF, bioA, bioD* and *bioB* in all biotin-synthesizing organisms known, including *Bacillus subtilis* (83, 110) (Figure 1-2). This section of the pathway begins with the incorporation of a pimelate (a seven carbon α, ω-dicarboxylic acid) thioester to eventually generate the heterocyclic rings of biotin in four enzymatic steps. One of the first demonstrations of biochemical activity was shown when pimeloyl-CoA and L-alanine were successfully converted to 8-amino-7-oxononanoic acid by *E. coli* cell free extracts, now known to be the function of BioF (29). Since then, *E.coli* has been used as a model organism to characterize all four enzymes that share sequence, structure and function to a high degree across organisms.

The first step of the second stage is catalyzed by BioF, a 8-amino-7-oxononanoate synthase. It catalyzes the formation of 8-amino-7-oxononanoate (also known as 7-keto-8-aminopelargonic acid, KAPA) from pimelate thioester. The enzymatic activity of BioF has been studied extensively in several organisms including *Bacillus sphaericus* (60, 106, 105), and the crystal structures of BioF from *E. coli* and *Mycobacterium smegmatis* are

available (2, 34). BioF belongs to a subclass II aminotransferase family of pyridoxal-5`-phosphate (PLP) dependent enzymes that catalyze condensation between an amino acid and a carboxylic acid CoA thioester of resulting in decarboxylation of the amino acid. It is a homodimeric protein with each dimer consisting of three domains — N-terminal, central and C-terminal — shaped like an open left hand (2). PLP is bound to a conserved lysine residue that plays a role in binding the substrate l-alanine, while phosphate groups of pimeloyl-CoA are thought to bind to the region of enzyme rich in positive arginine residues (2). The active site residues are conserved within the family members which share the reaction mechanism that involves Schiff base chemistry (2).

BioA is a 7, 8-diamino-nonanoate synthase that catalyzes the transformation of 8-amino-7-oxononanoate to 7, 8-diamino-nonanoate (also known as 7, 8-diaminopelargonic acid, DAPA). It also belongs to PLP-dependent enzyme family of the subclass III aminotransferases (65). The enzymatic activity was first characterized in cell-free extracts of *E. coli* (101) and then in other organisms such as *Mycobacterium tuberculosis* (23) and *B. subtilis* (124). S-Adenosyl-L-methionine (SAM), a common methyl donor, is the amino donor for *E. coli, B. sphaericus* (59) and *M. tuberculosis* BioA which is the only enzyme known to use SAM as the amino donor. *B. subtilis* BioA, on the other hand, uses L-lysine as an amino donor (124). Van Arsdell and co-workers argued that the preference of BioA for SAM or lysine corresponds to the intracellular level of the amino donor in the cells. They speculated that *E. coli* and *B. subtilis* can maintain a limited flux through BioA reaction by having a $K_m$ value for the amino donor at equal or higher concentration than the intracellular pool, regardless of the type of amino donor (124).

BioD is a dethiobiotin synthetase that catalyzes the ATP-dependent carboxylation of 7, 8-diamino-nonanoate to form dethiobiotin. Enzymatic and structural analysis of dethiobiotin synthase has been done in *E. coli* (69), *M. tuberculosis* (23) and *Helicobacter pylori* (108). This homodimeric enzyme belongs to the family of ATP-dependent carboxylases that contains the classical mononucleotide-binding motif common to all ATP and GTP-binding proteins (53). BioD represents the third mechanism

of carboxylation reaction other than those carried out by biotin dependent carboxylases and ribulose bisphosphate carboxylases (54).

The final step of biotin synthesis is catalyzed by BioB, biotin synthase. The reaction involves the insertion of a sulfur atom between the unactivated methyl and methylene carbon atoms adjacent to the ureido ring of dethiobiotin to form biotin. BioB is an iron-sulfur cluster containing homodimeric enzyme belonging to the radical SAM superfamily (7). Detailed mechanism involves a complex SAM-dependent radical chemistry that results in abstraction of the sulfur atom from its $[2Fe\text{-}2S]^{2+}$ cluster (122). BioB was also termed a "suicide enzyme" because removal of sulfur atom from $[2Fe\text{-}2S]^{2+}$ cluster renders it inactive for the next round of catalysis (7). Since the cluster is buried deep in the enzyme (7), rebuilding it requires the enzyme to unfold, which exposes it to proteolytic degradation (14). Biotin synthase has been shown to be catalytically active both *in vivo* (15) and *in vitro* (35) quelling previous doubts of BioB being a reactant rather than an enzyme. Despite low turnovers of biotin synthase, biological processes can continue due to low requirement for biotin (100 molecules/cell in *E. coli*) (14).

Diversity in first stage of synthesis

The first stage of biotin synthesis was recently elucidated in *E.coli* by Lin and coworkers (1, 76, 75) after a long history of ambiguity concerning the genes and the intermediates involved in the pathway. In contrast to the second stage of the pathway, there is no conserved way to generate the pimelate thioester intermediate in microbes due to the range of gene diversity (Figure 1-5). Of the relatively well studied organisms, *E. coli* and *B. subtilis* have different sets of genes predicted for pimelate thioester generation: *bioC* and *bioH* in *E. coli* (6, 111), and *bioW* and *bioI* in *B. subtilis* (10). Several other organisms either have the same genes or *bioC* with *bioH* homologs, *bioG* in *Haemophilus influenza* and *Campylobacter jejuni* (110), *bioK* in *Prochlorococcus marinus* (117), *bioJ* in *Francisella novicida* (37) or *bioV* in *Helicobacter sp.* (8). On the other hand, close relatives of *B. subtilis* (*B. licheniformis, B. amyloliquefaciens*) possess both *bioW* and *bioI.* Two genes, *bioW* and *bioX*, are identified in *Bacillus* (now

*Lysinibacillus*) *sphaericus* and *Staphylococcus aureus* whereas *Methanococcus janaschii* and *Chlamydia trachomatis* have a lone *bioW* (110). Another unique gene *bioZ* has been identified on the acquired symbiosis island of *Mesorhizobium loti* (12). Hence, it is possible to have multiple mechanisms at work in diverse organisms for the first stage of pimelate thioester intermediate synthesis.

The idea of pimelate being a precursor of biotin originated after Mueller, duVigneaud and coworkers described its growth stimulating effect on cultures of *Corynebacterium diphtheria* in biotin-free media (24, 88). Twenty years later, multiple researchers demonstrated the incorporation of $^{14}$C-pimelic acid into biotin vitamers in *Phycomyces blakesleeanus* and *Aspergillus niger* providing the first direct evidence of pimelic acid as a precursor for biotin (31, 32). Subsequently, transformation of pimeloyl-CoA to KAPA was demonstrated in cell-free extracts of *E. coli* (29), *B. sphaericus* (60) and several other microbes (57), providing a strong evidence for pimeloyl-CoA to be a likely intermediate in *E. coli* biotin synthesis. However, pimelic acid had no effect on *E. coli* biotin synthesis, which was speculated to be due to lack of permeability, as shown by the inability of *E. coli* to uptake $^{14}$C-pimelic acid (102).

Concomitant NMR studies of $^{13}$C-labeled biotin originating from [1-$^{13}$C] acetate or [2-$^{13}$C] acetate established a model for a plausible pathway for pimeloyl-CoA synthesis in *E. coli* similar to fatty acid synthesis and polyketide synthesis (55, 113). Results from these studies provided evidence for a pathway previously described by Lezius and coworkers (74). In this pathway, the seven-carbon dicarboxylate is synthesized by condensing three malonyl-CoA precursors while retaining the ω-carboxyl group of the initial malonyl-CoA. Additionally, these experiments dictated the origin of all the carbon atoms of biotin and provided concrete evidence to disqualify free pimelic acid as a precursor for *E. coli* biotin synthesis. The unique labeling pattern of biotin obtained from [1-$^{13}$C] acetate or [2-$^{13}$C] acetate indicate that the two carboxyl carbons of pimelic acid are metabolically distinct (Figure 1-6). Hence, pimelic acid could not be a free molecule and had to be attached to a carrier molecule (CoA, acyl carrier protein, condensing enzyme) via thioester linkage before its incorporation into biotin. Sanyal and

coworkers also noted that the labeling pattern eliminated other probable pathways for pimelate/pimeloyl-CoA synthesis via oxidation of octanoate, elongation of 2-ketoglutarate, or derivation from the tryptophan, lysine or diaminopimelate pathways (113). However, no concrete pathway could be determined for decades because studies failed to elucidate the precise function of *bioC* and *bioH* in *E. coli.*

**The BioC-BioH pathway of *Escherichia coli***

It was not until 2010 when Lin and coworkers made a breakthrough on the detailed mechanistic roles of BioC and BioH to generate pimelate moiety by utilizing the existing fatty acid synthesis pathway (Figure 1-3). It was revealed through insightful experiments that BioC transfers a methyl group from SAM to malonyl-ACP to form malonyl-ACP methyl ester (76). The methyl group serves a critical purpose to protect the charged carboxyl end of malonyl moiety so it can be tolerated by hydrophobic active sites of fatty acid enzymes for two rounds of the normal reduction, dehydration, reduction cycle to generate pimeloyl-ACP methyl ester (75). BioC was found to prefer malonyl-ACP over malonyl-CoA for the methyl acceptor, which changed our understanding about the acyl-carrier molecule for biotin synthesis in *E. coli.* Acyl carrier protein (ACP) and coenzyme A (CoA) share the phosphopantetheine group. It is esterified to an adenosine mono-phosphate in CoA, and in ACP the hydroxyl group of a serine residue of the protein is linked to the pantotheine moiety through a phosphodiester linkage (125). Being similar in structure, they can often be used interchangeably in *in vitro* enzyme assays that require either CoA or ACP but with different affinities (77). BioH then cleaves the methyl ester bond to generate pimeloyl-ACP, preventing further chain elongation (1). Hence, contrary to previous idea of pimeloyl-CoA being the intermediate, pimeloyl-ACP was revealed to be physiologically relevant intermediate in *E. coli* biotin synthesis. Organisms with the same genes or its homologs are expected to have an analogous pathway for the first stage of biotin synthesis.

**The BioW-BioI pathway of *Bacillus subtilis***

The absence of genes homologous to *bioC* and *bioH* in *B. subtilis* and many other organisms indicates a departure from the above mentioned pathway for pimelate moiety

synthesis. Comparative genomic analyses revealed *bioW* as the more frequent replacement for *bioC* and *bioH* as opposed to *bioZ*, while *bioI* or *bioX* seem to accompany *bioW* in some of the genomes (110). BioW is annotated as a pimeloyl-CoA synthetase that catalyzes ATP-dependent transfer of pimelic acid to CoA, resulting in pimeloyl-CoA. Izumi and coworkers first demonstrated the formation of pimeloyl-CoA from pimelic acid, ATP, CoA and $Mg^{2+}$ in *Bacillus megaterium* (61, 58), naming the enzyme pimeloyl-CoA synthetase. Gloeckler and coworkers characterized the *bio* genes of *B. sphaericus* that result in dethiobiotin from pimelate (40). Genetic complementation studies revealed the ability of *bioW* to bypass the requirement for pimelate thioester intermediate conferred by Δ*bioC* Δ*bioH* mutations in *E. coli*, when pimelic acid is supplied in the minimal medium. Subsequently, the enzymatic activity of *B. sphaericus* BioW was characterized and confirmed as a pimeloyl-CoA synthetase (107). Hence, *bioW* requiring organisms like *B. subtilis* appear to have a direct mechanism to utilize free pimelic acid in a one-enzyme reaction dedicated to pimelate thioester synthesis.

*B. subtilis* represents a group of microbes that harbor two genes for the same function, synthesis of pimelate thioester moiety. Bower and coworkers extensively characterized *B. subtilis bio* operon with *bioW* and *bioI* in a single unidirectional operon along with *bioA*, *bioF*, *bioD* and *bioB* (10). Genetic deletion and complementation studies revealed both *bioW* and *bioI* are independently capable of bypassing the requirement for pimeloyl-ACP in *E. coli ΔbioC* and *ΔbioH* mutant strain. Genetic mutation analysis revealed that *bioW* is important for biotin synthesis since its mutation caused biotin auxotrophy, while *bioI* causes a bradytrophic phenotype (10). Biochemical data on BioI argues otherwise, suggesting that BioI is the primary enzyme responsible for the generation of pimeloyl-ACP. *B. subtilis* BioI has been extensively characterized compared to BioW. BioI is a cytochrome P450 with significant sequence homology and spectral properties characteristic of other P450s that show a spectrophotometric peak at 450 nm when bound to carbon-monoxide in a reduced state (39, 43, 100, 119). BioI is one of the eight cytochrome P450s revealed to be present in *B. subtilis* genome (70). Cytochrome P450s constitute a large family of hemoproteins that are capable of catalyzing oxidation of a vast range of organic compounds (73, 92). The mechanism of

C-C bond cleavage involves activation of molecular oxygen to produce a highly reactive Fe (V)-oxo species from two rounds of NAD(P)H dependent flavodoxin/ferredoxin-mediated reductions (92). *B. subtilis* Fer protein is believed to be the electron donor ferredoxin, however, the partner ferredoxin reductase has not been found (42). *orf2*, a gene downstream to *bioI*, has sequence similarities to putative NADH- or NADPH-binding sites, however, deletion of *orf2* had no effect on biotin synthesis (10).

Stok and coworkers purified heterologously expressed BioI bound to *E. coli* acyl-ACP ranging in length from C14-C18 (119). Liberation of fatty acids from BioI-ACP complex under basic conditions yielded pimelic acid, albiet to a low percent, inciting the authors to predict BioI as a source of pimeloyl-ACP for biotin synthesis. This finding supported previous genetics reports of *bioI* mutants causing a bradytrophic phenotype (10). Turnover of free fatty acids to low amounts of pimelic acid by BioI was also detected. Green and coworkers demonstrated a mechanism of cleaving free fatty acid via ω-hydroxylation followed by chain shortening reactions to produce pimelic acid (43). Later, high resolution crystal structure of BioI-acyl-ACP complex revealed a hydrophobic binding pocket above the active-site heme iron, where the incoming fatty acid chain adopts a kinked U-shaped conformation (19). The precise placement of C7 and C8 atoms above the heme iron independent of fatty acid chain length clearly indicated C-C cleavage that results in pimeloyl-ACP by three rounds of oxidations (119). The specificity of BioI catalytic activity was attributed to ACP protein being bound to the fatty acids in absence of which an assortment of products of different chain lengths are generated (19, 119). Hence, biochemical studies suggest that BioI contributes to the biotin biosynthetic pathway by generating pimeloyl-ACP whereas pimelic acid might be a byproduct of BioI catalytic activity, which is yet to be examined.

*B. subtilis* therefore seemed to have a redundant system in the first stage of the pathway (Figure 1-4). It needed to be determined whether BioI and BioW function in conjunction or independently. It is logical to postulate that BioI can generate a pimelate thioester moiety independent of BioW to produce biotin, serving as a salvage pathway rather than an obligatory pathway. Since deletion of *bioW* conferred a more deleterious

effect on growth, BioW could be the essential enzyme involved in the generation of pimelate thioester from pimelic acid possibly via BioI or some other unknown source. The pathway for synthesis of pimelic acid is undetermined. Previous studies have reported the formation of pimelic acid as a by-product of azelaic acid degradation in yeast (97) *Pseudomonas sp.* (62) and *Micrococcous sp.* (95). Pimelic acid has also been shown to be synthesized from $C_{18}$ fatty acids such as exogenously supplemented oleic acid (18:1 *cis*-9), thought to be the precursor of azelaic acid (nine carbon α, ω-dicarboxylic acid), and linoleic acid (18:2 *cis*, *cis*-9,12) in various bacteria (98, 99). However, unlike most living organisms that have straight-chain acids with or without unsaturation in the carbon chain, *B. subtilis* has major amounts of lipids made of rare branched-chain fatty acids having a methyl group at the iso or anteiso positions (66). This adds to the conundrum for the synthesis of pimelate, a seven-carbon straight chain dicarboxylate.

The redundancy in pimelate thioester generation also raises an important question about the substrate specificity of BioF, the next enzyme that utilizes the pimelate thioester intermediate to make KAPA. Since there are two derivatives of pimelate moiety available, does it accept both or prefer one over the other for KAPA synthesis? *B. subtilis* BioF has 50% amino acid sequence identity to the BioF of *B. sphaericus* which has been well characterized for its synthesis of KAPA from pimeloyl-CoA (106, 105). *E. coli* BioF, however, was reported to have a lower affinity for pimeloyl-CoA compared to that of *B. sphaericus* (128). Since pimeloyl-ACP is the intermediate in *E. coli,* I hypothesize *E. coli* BioF to have higher affinity with pimeloyl-ACP compared to the CoA moiety which is the only pimelate derivative tested so far. Hence, *E. coli* and *B. subtilis* BioF may be diverse in their specificity for different pimelate thioester moieties because of the difference in how the intermediates are synthesized. Therefore, determining the substrate specificity of *B. subtilis* BioF will help elucidate the physiologically relevant carrier molecule of pimelate and also help resolve the question about the redundancy of BioW/BioI pathway.

The existence of two pathways to generate pimelate thioester in *B. subtilis* is unusual. Speculations can be made about the reason as to why the organism has two

enzymes, possibly for the same purpose. I hypothesize that BioW is the important enzyme to generate pimeloyl-CoA while BioI is dispensable for biotin synthesis. This is based on the evidence that the gene *bioI* has not been found to exist in an operon without *bioW.* Only *B. subtilis* and a few close relatives have genes that are identified as *bioI.* However, *bioW* has been found in a diverse range of organisms as a sole provider of pimelate thioester moiety (110). Furthermore, *bioI* is the last gene in the *bio* operon with a preceding putative terminator and is transcribed 8-fold lower than the rest of the operon upstream to it (104). BioI might be an evolutionary remnant when *B. subtilis* and its close relatives were strict aerobes depending on a cytochrome P450 enzyme for the synthesis of biotin. *B. subtilis* is a soil bacterium and it might have acquired the new gene *bioW* to learn to adapt to different oxygen levels present in soil hence, increasing its chances of survival. Indeed, *B. subtilis* was regarded as a strict aerobe until a review by Priest mentioned its ability to grow slowly under strictly anaerobic conditions with nitrate as terminal electron acceptor (93, 109).

## SPECIFIC AIMS OF THIS THESIS

I started to work on this Thesis project to answer three major questions about biotin synthesis in *B. subtilis*: the redundancy of pimelate thioester generation, origin of pimelic acid and the substrate specificity of BioF. In the following chapters, I have described the approaches taken to answer these long-standing questions.

In Chapter 2, I characterized the enzymatic properties of BioW to provide evidence for its pimeloyl-CoA synthetase function. BioW is a unique enzyme belonging to a pimeloyl-CoA synthetase family that is different from any other acyl-CoA synthetase known in terms of sequence and structure. I describe a unique characteristic of *B. subtilis* BioW to proofread non-cognate acyl-adenylate intermediates, a characteristic similar to the editing function of amino-acyl tRNA synthetases. The reaction of pimeloyl-CoA synthesis is a two-step process that involves the formation of an acyl-adenylate intermediate, which is transformed into acyl-CoA. I found that BioW is able to proofread glutaryl- (five carbon) and suberyl-(eight carbon) adenylate and thus, prevents the

synthesis of shorter and longer chain acyl-CoA for biotin synthesis. Recently, a collaboration with Dr. Satish Nair resulted in the first crystal structure of BioW from *Aquifex aeolicus* which provided insights on what residues might be involved in the binding and proofreading functions of the enzyme.

In Chapter 3, I demonstrate the significance of BioW as the sole producer of pimelate thioester intermediate for *B. subtilis* biotin synthesis. BioI is capable of generating pimeloyl-ACP. However, this function seems to be wasteful in *B. subtilis* since absence of *bioI* (gene mutation) and growth in anaerobic condition (inactive *bioI*) does not affect biotin synthesis in minimal media. Moreover, I made advances toward elucidating the source of pimelic acid *in vivo.* In $^{13}$C-NMR experiments similar to Sanyal and coworkers, I provide the first evidence for pimelic acid being a *bona fide* precursor of biotin. I discovered evidence for pimelate synthesis via head-to-tail chain elongation process similar to fatty acid synthesis, which is supported by the negative effects of fatty acid inhibitors on biotin synthesis *in vivo.*

In Chapter 4, I summarize all my findings to provide a better understanding about the biotin synthetic pathway in *B. subtilis* and other organisms that are pertinent to this study.

FIGURES

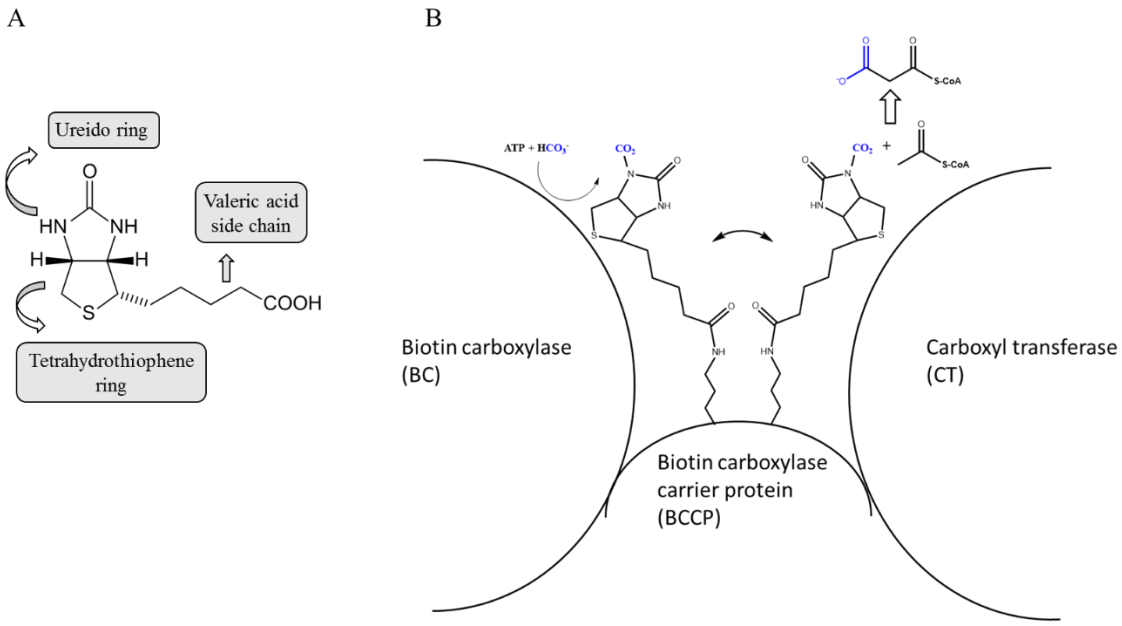



**Figure 1.1:** A) Structure of biotin. It consists of an ureido ring and a tetrahydrothiophene ring attached to a valeric acid side chain. B) Function of biotin. Biotin attaches to the BCCP subunit of biotin dependent enzymes forming an amide bond with a conserved lysine residue. Biotin is carboxylated in an ATP-dependent manner by biotin carboxylase using bicarbonate as a source of $CO_2$. It acts as a swinging arm and travels to the carboxyl transferase subunit, where the carboxyl group is transferred from biotin to acetyl-CoA to generate malonyl-CoA, in case of acetyl-CoA carboxylase.

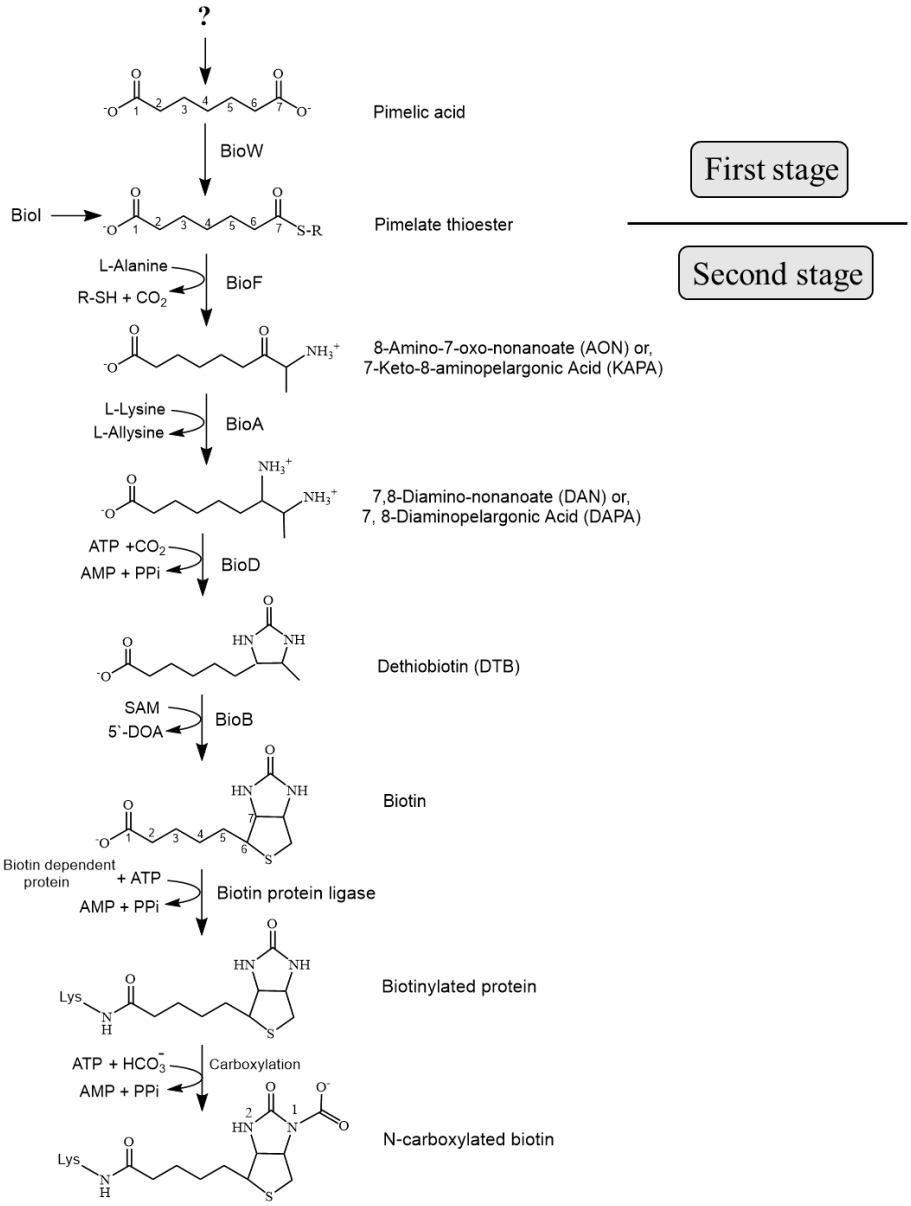



**Figure 1.2:** The biotin biosynthesis pathway in *B. subtilis.* It can be separated into two distinct stages. First stage represents two possible pathways to synthesize pimelate thioester intermediate via either BioW (pimeloyl-CoA synthetase) or BioI (a cytochrome P450). Pimelate thioester may be in the form of pimeloyl-CoA or pimeloyl-ACP. The second stage of the pathway involves BioF driven synthesis of 8-amino-7-oxo-nonanoate (AON) which is converted to 7, 8-diamino-nonanoate (DAN) by BioA. BioD transforms DAN into dethiobiotin (DTB) which is finally converted to biotin by BioB. Biotin is attached to a conserved lysine residue of a biotin dependent enzyme in an ATP dependent manner. After attachment, biotin is carboxylated on N1. Abbreviations: SAM, *S*-adenosyl-L-methionine; 5'-DOA, 5'-deoxyadenosine; ATP, Adenosine -5`-triphosphate. This figure is adapted from a review by Lin *et al* (77).

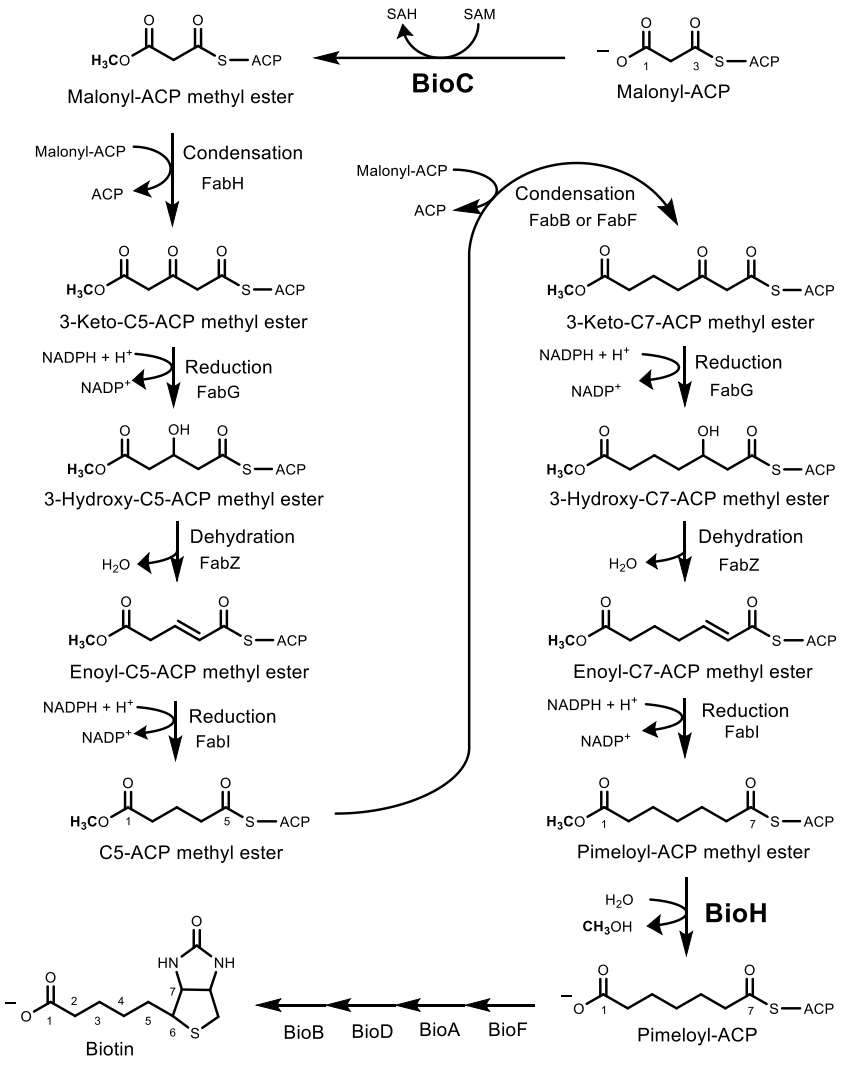



**Figure 1.3:** The initial stage of biotin synthesis in *E. coli* via BioC and BioH. BioC catalyzes the SAM-dependent transfer of methyl-group to malonyl-ACP to produce malonyl-ACP methyl ester. The chain elongated from three carbons of malonyl moiety into seven carbons of pimeloyl moiety in two rounds of reduction, dehydration and reduction by fatty acid enzymes to generate pimeloyl-ACP methyl ester. At this point, further chain elongation is halted by BioH by removing the methyl-group from the pimeloyl moiety. Pimeloyl-ACP is then converted to biotin by the second stage of the pathway that involves BioF, BioA, BioD and BioB. Abbreviation: SAH, *S*-adenosylhomocysteine. This figure is adapted from Lin *et al* (75).

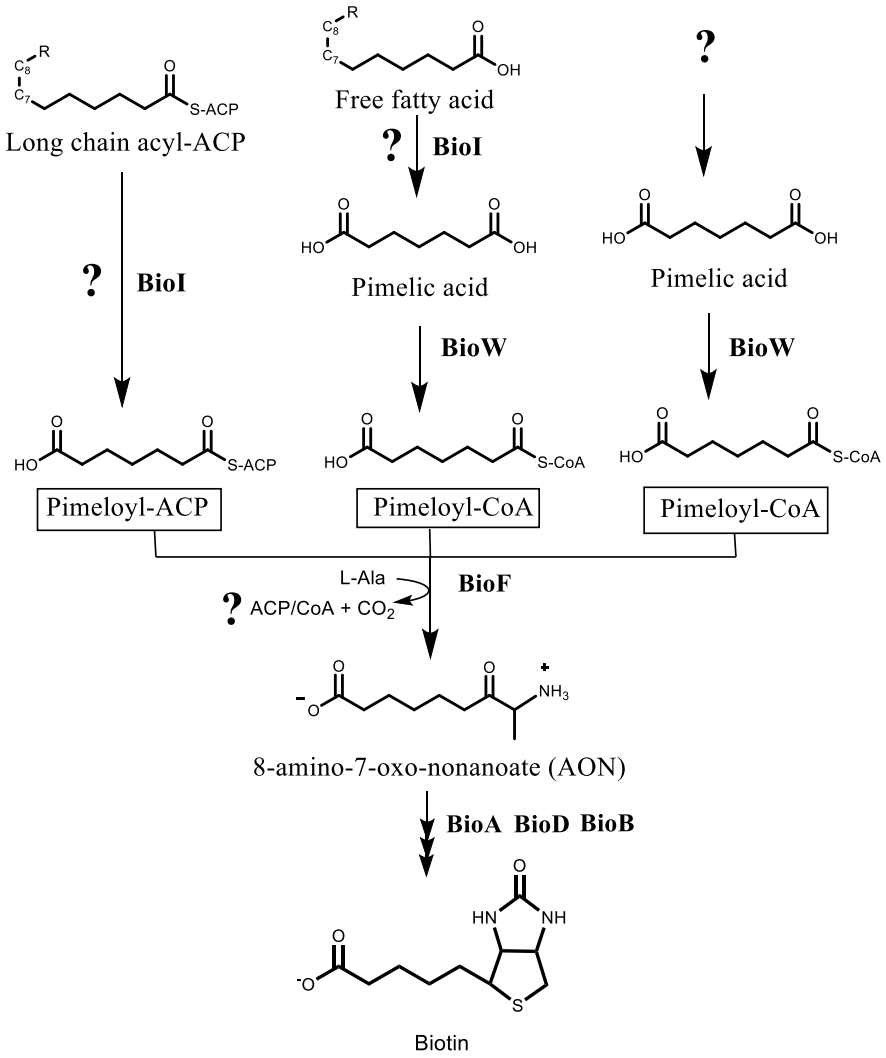



**Figure 1.4:** Schematic representation of three possible pathways for pimelate thioester synthesis via BioI and BioW. In the first pathway, BioI generates pimeloyl-ACP from C-C bond cleavage of long chain acyl-ACP. In the second pathway, BioI provides pimelic acid via C-C bond cleavage of free fatty acids and BioW synthesized pimeloyl-CoA. In the third pathway, BioW acquires pimelic acid from a yet unknown synthesis pathway(s) to produce pimeloyl-CoA. Finally, it is upon BioF to decide which intermediate to use, pimeloyl-ACP or pimeloyl-CoA for the completion of biotin synthesis.

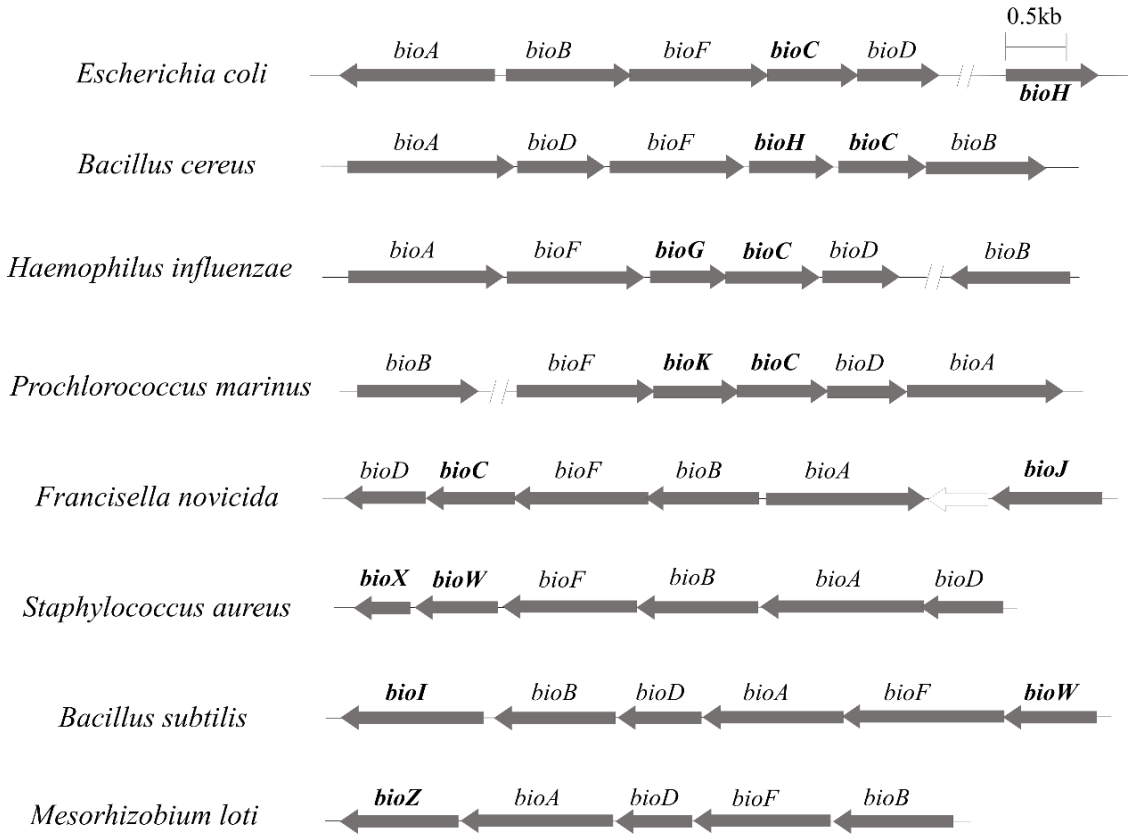



**Figure 1.5:** Diversity of gene arrangement in *bio* operon. The genes involved in the initial stage of biotin synthesis are represented in bold font. Diverse organisms like *E. coli* and *B. cereus* contain both bioC and bioH. *bioH* is replaced by its homologs *bioG*, *bioK* and *bioJ* in *H. influenzae, P. marinus* and *F. novicida* respectively. *S. aureus* and *B. subtilis* contain *bioW* and *bioX* or *bioI* that share no sequence homology to *bioC* or *bioH.* Rhizobia like *M. loti* contain *bioZ* which differs from both *bioW* and *bioC/bioH.*

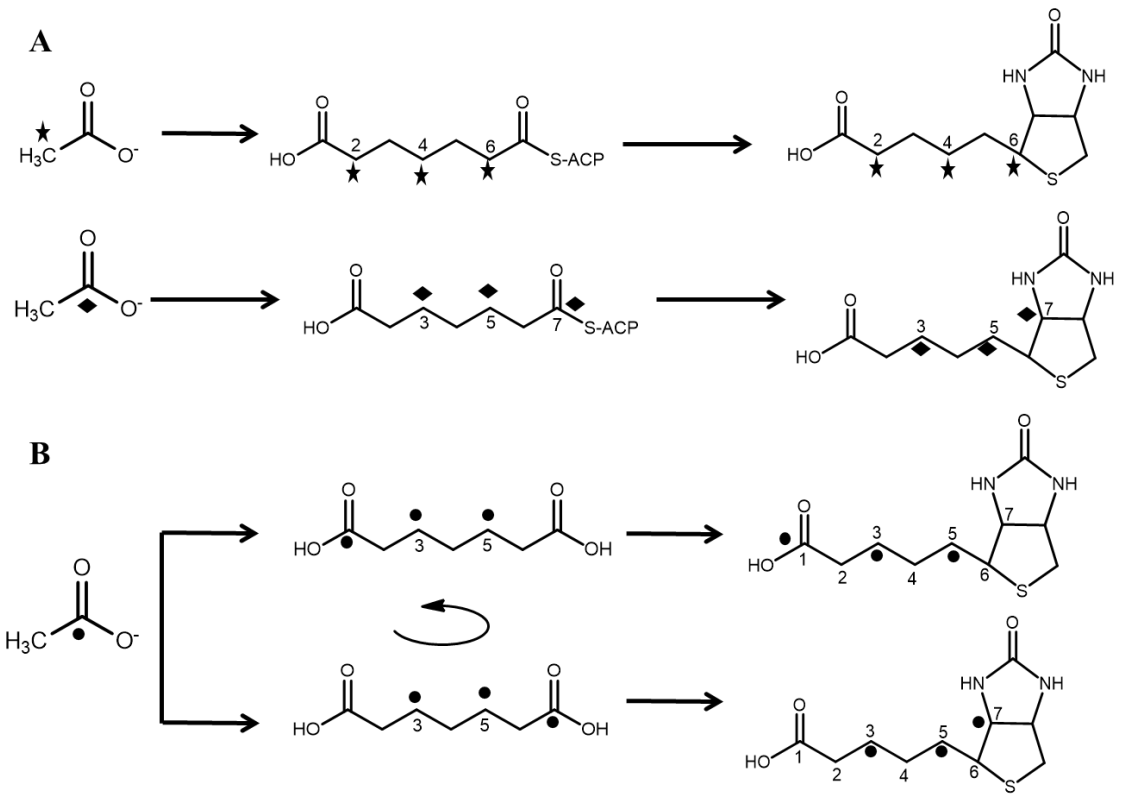



**Figure 1.6:** Biotin labeling patterns originating from [1-$^{13}$C] or [2-$^{13}$C] acetate. A) Schematic representation of 13C-NMR results obtained by Sanyal et al and Ifuku et al in *E. coli*. Carbons 2,4 and 6 were labeled in biotin originating from [2-$^{13}$C] acetate while carbons 3,5 and 7 were labeled in biotin originating from [1-$^{13}$C] acetate. This labeling pattern requires pimelate moiety to be attached at one end to a carrier molecule. B) Hypothetical representation of $^{13}$C-labeled biotin in *B. subtilis* to show that pimelic acid is a *bona fide* precursor of biotin synthesis. Unlike in *E. coli*, the symmetrical pimelic acid molecule can be incorporated in either direction to have both the carboxyl terminals of biotin labeled to the same extent.

# CHAPTER 2
# PROOFREADING OF NONCOGNATE ACYL ADENYLATES BY AN ACYL-COENZYME A LIGASE

## INTRODUCTION

A major determinant of the low error rate of protein synthesis is the specific aminoacyl-tRNA synthetase-catalyzed acylation of tRNAs by only the cognate amino acid. Aminoacyl-tRNA synthetases can ensure faithful protein synthesis by having high substrate selectivity. However, as first pointed out by Pauling (103), small differences in binding energy between aliphatic amino acids cannot provide the discrimination necessary for faithful protein synthesis. For amino acids having closely related structures such as isoleucine and valine, discrimination against the smaller amino acid would be incomplete. Indeed, Baldwin and Berg (4) showed that isoleucine tRNA synthetase converted both isoleucine and valine to their adenylates. However, only isoleucine was transferred to tRNA because valyl-adenylate was hydrolyzed to AMP and valine. These acyl-adenylate editing reactions proceed either by eviction of the non-cognate acyl-adenylate into solution where it undergoes spontaneous hydrolysis or by catalyzed hydrolysis within the active site (or in a separate editing site). In some cases enzymatic hydrolysis is stimulated by the presence of an acceptor tRNA. These reactions are generally referred to as editing, but proofreading is more accurate because proofreading only eliminates errors whereas editing improves the product.

I report proofreading by an acyl-Coenzyme A (CoA) synthetase/ligase (synthetase and ligase are approved synonyms). The acyl-CoA ligase is BioW, an enzyme required for biotin synthesis in *Bacillus subtilis* and closely related bacteria (10). The physiological reaction of BioW is the ATP-dependent conversion of pimelate, a seven carbon α, ω-dicarboxylic acid, to its CoA monothioester (Figure 2-1A). Although BioW contains none of the sequence motifs characteristic of acyl-CoA ligases, the reaction was demonstrated to proceed through the canonical acyl-adenylate intermediate. However, when presented with glutaric acid, the C5 homologue of pimelic acid, only traces of

adenylate and CoA thioester were formed, and much of the ATP was converted to AMP. Conversion of ATP to AMP in the absence of final product synthesis is the hallmark of pretransfer editing in protein synthesis (130). My data argue that most of the glutaryl-adenylate intermediate is largely cleaved within the BioW active site. A similar, but less efficient, proofreading of the adenylate of a second incorrect substrate, the C8 α, ω-dicarboxylic acid, suberic acid, also occurred.

Pimelic acid (heptanedioic acid) is a seven-carbon α, ω-dicarboxylic acid that contributes most of the carbon atoms of biotin, the others coming from alanine and $CO_2$ (77) (Figure 2-1A). To be used as a biotin precursor one of the pimelate carboxyl groups must be activated by conversion to a thioester (of either CoA or acyl carrier protein) (Figure 2-1A). This thioester then reacts with alanine in the decarboxylative condensation reaction catalyzed by BioF to form 8-amino-7-oxononanoate, the first intermediate in formation of the fused heterocyclic rings of biotin (Figure 2-1A).

## MATERIALS AND METHODS

### Bacterial strains and plasmids

The *bioW* gene was PCR amplified from *Bacillus subtilis* subsp. subtilis str. 168 genomic DNA using primers pMM1 5`-GGA GAT ATA CCA TGA TGC AAG AAG AAA C-3` and pMM2 5`-CTT CTT GCA TCA TGG TAT ATC TCC TTC-3` and Pfu Turbo DNA polymerase from Stratagene. The amplified gene was digested with restriction enzymes, NcoI and XhoI, and inserted between the NcoI and XhoI sites of vector pET28b (Novagen). The sequence verified plasmid (pMM8) was transformed into *E. coli* BL21 Tuner (Novagen) and the transformants were plated on LB medium (per liter, 10 g of tryptone, 5 g of yeast extract, and 5 g of NaCl) containing 50 μg/ml of kanamycin sulfate.

### Expression and purification of BioW

A 2 L LB culture of *E. coli* BL21 Tuner transformed with plasmid pMM8 was grown to an optical density 600 of 0.6 at 37°C followed by induction with 1 mM

isopropyl-β-D-thiogalactopyranoside for 5 hr at 37°C. The cells were harvested by centrifugation and resuspended in 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10% glycerol and 5 mM tris (2-carboxyethyl) phosphine hydrochloride (TCEP), and then they were lysed by passage through a French pressure cell. The supernatant obtained after centrifugation of cell lysate at 39,000x*g* for 30 min at 4°C, was applied to a 5 ml HiTrap Q FF ion exchange chromatography column (GE Healthcare) equilibrated with the buffer used to suspend the harvested cells. The column was washed with 50 mM Tris-HCl (pH 7.5), 1 M NaCl, 10% glycerol and 5 mM TCEP. The protein fraction eluted at about 250 mM NaCl and was analyzed by SDS/PAGE. Fractions containing the highly purified protein were pooled, concentrated and resuspended in 50 mM Tris (pH 7.5), 100 mM NaCl, 10% glycerol, 5 mM TCEP and 1 M ammonium sulfate and applied to a 5 mL HiTrap Phenyl HP hydrophobic interaction chromatography column (GE Healthcare) equilibrated with the same buffer. The column was washed with 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10% glycerol and 5 mM TCEP, and the protein was eluted with a linear gradient from 1 M to 0 ammonium sulfate. After analyzing samples by SDS/PAGE, the fractions having the highest degree of purity were pooled and applied to a HiLoad 26/60 Superdex 200 size exclusion chromatography column. The column was equilibrated with 50 mM Tris-HCl (pH 8.0) containing 150 mM NaCl and the protein was allowed to elute over one column volume. The protein fractions of greatest purity were pooled and dialyzed against a storage buffer of 50 mM Tris-HCl (pH 8), 150 mM NaCl, 20% glycerol and 5 mM TCEP. Finally, the dialyzed protein was concentrated to 2.7 mg/mL, flash frozen and stored at -80°C.

High Performance Liquid Chromatography (HPLC) enzyme assays

Pimeloyl-CoA ligase activity was assayed by detection of pimeloyl-CoA formation. The reactions contained 100 mM sodium HEPES (pH 7) buffer, 10 mM $MgCl_2$, 0.1M NaCl, 0.2 mM TCEP, 5 mM pimelate, 1 mM CoA and 1 mM ATP and were initiated by adding enzyme to 5 µM (all enzyme concentrations are given as monomeric protein). After incubation at 37°C for 1 h, the reaction was stopped by adding 10% trichloroacetic acid to precipitate the protein. The precipitated protein sample was removed by centrifugation and filtration through Amicon Ultra 3K devices (Millipore).

The filtrate was loaded onto a µBondapak C18 cartridge (Waters) previously equilibrated with 50 mM ammonium acetate (pH 5). Solvent A was 50 mM ammonium acetate (pH 5) and solvent B was acetonitrile. The sample was eluted with 0% B for 10 min, 30 min of a gradient from 0%B to 50% B followed by 5 min gradient from 50% B to 80% B which was maintained for 10 min (56). Eluent corresponding to the pimeloyl-CoA peak was dried under nitrogen and the mass was analyzed at the University of Illinois, School of Chemical Sciences Mass Spectrometry Laboratory by electrospray ionization mass spectrometry performed on a Micromass Q-Tof Ultima instrument run in the positive mode.

Thin layer chromatographic enzyme assays

In general, the reactions contained 100 mM sodium HEPES (pH 7) buffer, 5-10 mM $MgCl_2$, 0.1-0.2M NaCl, 0.2 mM TCEP, 1-10 mM pimelate, 0.25 mM CoA, 25-100 µM ATP with [α-$^{32}$P]ATP 600 to 1200-fold lower in concentration than the nonradioactive ATP and 5 µM BioW, except where indicated. The specificity of BioW was tested with malonic acid (C3), succinic acid (C4), glutaric acid (C5), adipic acid (C6), pimelic acid (C7), suberic acid (C8), azelaic acid (C9). To assay non-cognate substrate editing by BioW, the rate the AMP formation in the absence of CoA was monitored in a time dependent manner. At each time point 1 µl of the reaction was spotted onto a thin-layer plate. Kinetic parameters for pimelate and CoA were determined by varying the concentration (5-100 µM) of the respective substrate in a reaction containing 50 nM BioW in order to achieve a linear progression of enzymatic activity. Thin-layer chromatography (TLC) on cellulose plates was developed in isobutyric acid: ammonium hydroxide: water (66:1:33) (16). After drying, the thin layer chromatograms were exposed to a phosphor imaging plate for 1-17 hr and the results visualized using a Fujifilm FLA-3000 instrument. The chromatogram spots were quantitated by comparing them with a standard curve generated by serial dilution of known [α-$^{32}$P]ATP concentration for each plate. All reactions were repeated three times and the kinetic parameters were determined by fitting the data points to Lineweaver-Burk plots.

## RESULTS AND DISCUSSION

Although in *Escherichia coli* pimelate is formed by a modified fatty acid synthetic pathway (75), the source of pimelate in *B. subtilis* remains unknown. However, the role of BioW is clear; strains lacking BioW are defective in biotin synthesis (see Chapter 3). Moreover, upon expression of BioW in *E. coli*, pimelate supplementation allows bypass of mutations in pimelate synthesis (10, 75). BioW is an extremely unusual acyl-CoA ligase in that it lacks all of the well-characterized motifs of this highly conserved enzyme family (45) and is half the typical size. This remarkable divergence from the canonical enzymes led us to purify BioW and characterize the pimeloyl-CoA synthesis reaction in detail.

Attachment of hexahistidine tags to either end of BioW destroyed the ability of the protein to function in *E. coli*; thus, the enzyme was purified by conventional column chromatography steps to obtain preparations that gave a single band in denaturing gel electrophoresis and a single peak on size exclusion chromatography that indicated BioW is a dimeric protein (Figures 2-4 and 2-5). Pimeloyl-CoA synthesis was assayed by high-pressure liquid chromatography (HPLC) with detection by UV absorbance, and required pimelate, ATP, $Mg^{2+}$ and CoA (Fig. 1b). Pimeloyl-CoA formation, verified by mass spectroscopy (Figure 2-1D), was accompanied by production of AMP consistent with a pimeloyl-adenylate intermediate. Use of [$\alpha$-$^{32}$P]ATP allowed direct demonstration of the pimeloyl-adenylate intermediate that accumulated in the absence of the CoA acceptor and was converted to AMP upon addition of CoA (Figure 2-1C). The overall reaction had the stoichiometry expected of an acyl-CoA ligase; one pimeloyl-CoA molecule was formed per ATP hydrolyzed to AMP (45). The reaction was most rapid at pH 7. ATP was the sole active nucleotide triphosphate and a metal ion ($Mg^{2+} > Mn^{2+} > Co^{2+}$) was required. (Figure 2-6). In addition, the kinetic parameters for pimelate and CoA were determined in sodium 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer at pH 7 using the $\alpha$-$^{32}$P-labeled ATP assay (Figure 2-1B). These data showed that BioW is a notably poor catalyst. This property is characteristic of biotin synthetic enzymes (5, 77, 75) consistent with the extremely low physiological demand for biotin.

The BioW dicarboxylate substrate specificity was first determined in the overall reaction, and as expected, pimelate was the best substrate (Figures 2-1A & 2-2B). No synthesis of the CoA ester was seen with the C6 homolog adipic acid, whereas the C8 acid suberic acid was only weakly active, verified by ESI-MS (Figure 2-7), and the C5 acid glutaric acid gave only traces of the CoA ester (Figure 2-2B). To determine which partial reaction limited use of the noncognate substrates, $\alpha$-$^{32}$P-labeled ATP was used to assay conversion of the acids to their acyl-adenylates in the absence of CoA. Adipic (C6) and suberic (C8) acids synthesized detectable levels of their adenylates, whereas the C5 homolog glutaric acid formed only trace amounts of glutaryl-adenylate (Figure 2-2C). However, in the glutarate reactions AMP accumulated at the expense of ATP in a glutarate-dependent manner (Figure 2-2C). AMP accumulation in reactions with a noncognate substrate is the hallmark of pretransfer proofreading (editing) by aminoacyl-tRNA synthetases; the adenylate is synthesized and then cleaved to AMP (85, 130). For some aminoacyl-tRNA synthetases, proofreading is largely or totally dependent on the presence of the tRNA cognate to the synthetase (130). However, the presence of CoA had no detectable effect on glutaryl-adenylate hydrolysis by BioW (Figure 2-2C). Note that the low levels of glutaryl-CoA accumulation cannot be attributed to solvent hydrolysis of the thioester. Glutaryl-CoA is a well-studied enzyme substrate and is not an unusually labile acyl-CoA. Indeed, I monitored the thioester linkage by absorbance at 232 nm of a solution of glutaryl-CoA at 23$^{o}$C for 3 hr in assay buffer at pH 7.5 without detectable loss of absorbance.

The C8-adenylate also showed signs of hydrolysis to AMP, whereas the C6-adenylate accumulated but showed no hydrolysis in the absence of CoA (Figure 2-2C). Upon CoA addition AMP levels increased somewhat, but no adipoyl-CoA was formed (Figure 2-2B), suggesting that CoA increased solvent hydrolysis of the enzyme-bound C6-adenylate. The BioW reaction with adipic acid (C6) was not further studied because it is not a naturally occurring compound and thus physiologically irrelevant.

As reported for aminoacyl-tRNA synthetase pretransfer editing reactions (130), adenylate hydrolysis could occur either by cleavage within the BioW active site or by release of the adenylate to solution where the mixed anhydride bond would be cleaved by hydroxide ion. The rate of acyl-adenylate hydrolysis increases as pH values rise above neutrality (21, 64, 114); thus, hydrolysis of adenylates released to solution should increase with pH. However, in the C5- and C8-adenylates, increased solvent pH had no effect on hydrolysis of the adenylate (Figure 2-3A). The most straightforward interpretation of these data is that these adenylates are largely sequestered within the BioW active site. If so, then hydrolysis of the C5- and C8-adenylates must proceed within the enzyme active site. In contrast, the level of the C7-adenylate declined at elevated pH values with a concomitant increase in AMP, indicating that solvent could access this adenylate (Figure 2-3A). A greater accessibility of C7-adenylate bound within the active site seems reasonable because the CoA thiol must enter the active site for pimeloyl-CoA production. However, a pH-induced conformational change that releases the C7-adenylate to solution could also occur. The BioW-catalyzed hydrolysis of the C5- and C8-adenylates was linear with time, whereas pimeloyl-adenylate filled the BioW active site (one molecule of pimeloyl-AMP per BioW monomer) whereupon AMP production essentially ceased (Figure 2-3C).

BioW proofreading by cleavage of glutaryl-adenylate is reminiscent of valine activation by isoleucine-tRNA synthetase, a situation first addressed by Pauling (103) who calculated that the small differences in binding energy among aliphatic amino acids could not provide the discrimination necessary for accurate protein synthesis. He estimated that incorporation of isoleucine would be favored over valine by a factor of only 100-200. Direct binding measurements to isoleucyl-tRNA synthetase confirmed Pauling's prediction (78). However, the error rate in insertion of valine for isoleucine is 1 in 3,000 (79) due to proofreading (38). Pauling's argument also pertains to the BioW reaction with glutarate. Glutarate is smaller than pimelate, precluding steric occlusion, and the two additional methylene groups of pimelate would not provide sufficient binding energy to insure specificity. Moreover, glutaric acid occurs in nature and is particularly abundant in seeds and root vegetables. Thus, it seems likely that *B. subtilis*, a soil

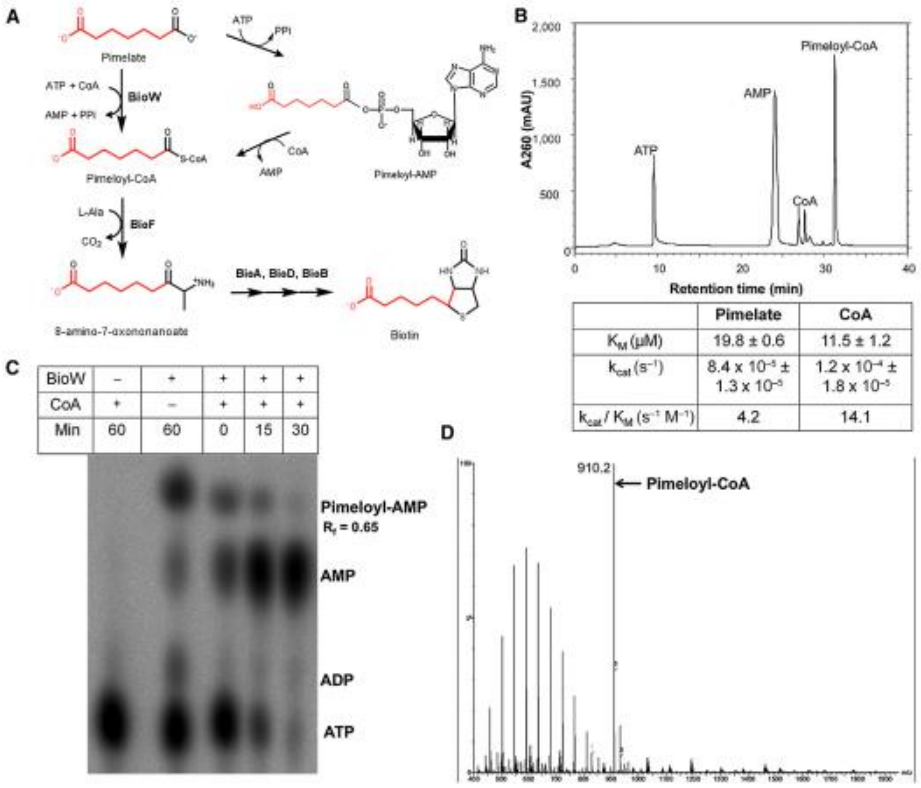
bacterium, could encounter high levels of this dicarboxylate and that proofreading of the adenylate would act to prevent possible short circuiting of biotin synthesis by glutarate.

Suberic acid also occurs naturally; thus, hydrolytic proofreading of its adenylate is appropriate. Superficially, isoleucine activation by valine-tRNA synthetase would seem an appropriate comparison, because the noncognate substrate has one additional methylene group. However, the chiral amino acid must bind the enzyme with strict stereochemistry, whereas the nonchiral and flexible dicarboxylate is able to bind in a manner that allows formation of the adenylate and low levels of the CoA ester.

The broad question is whether or not BioW proofreading has an evolutionary relationship with aminoacyl-tRNA synthetase pretransfer editing. This may be the case, but the extant protein sequences provide no support for a relationship. BioW is a most unusual protein in that database searches by a large number of search algorithms, including those designed to detect very remotely homologous sequences (e.g., HMMER http://hmmer.janelia.org/), give only BioW sequences. Moreover, as mentioned above, BioW has no recognized sequence motifs. This is not surprising for pimelate, an unusual ligand, but it is surprising for CoA and especially so for ATP. There may be structural homologies that could speak for (or against) a relationship between BioW and aminoacyl-tRNA synthetases, but this will require a BioW crystal structure.

Other examples of hydrolytic proofreading are known, although they involve different chemistry. For example, in polyketide synthesis a report demonstrated that a protein previously thought to be an acyltransferase was a thioesterase that specifically cleaved noncognate acyl-thioesters (64). It seems reasonable that proofreading reactions may occur most often in the early steps in synthesis of complex molecules such as biotin or an intricate polyketide to prevent embarking on the expensive synthesis of inactive molecules.

FIGURES




| | Pimelate | CoA |
|---|---|---|
| $K_M$ (μM) | 19.8 ± 0.6 | 11.5 ± 1.2 |
| $k_{cat}$ ($s^{-1}$) | 8.4 x $10^{-5}$ ± 1.3 x $10^{-5}$ | 1.2 x $10^{-4}$ ± 1.8 x $10^{-5}$ |
| $k_{cat}$ / $K_M$ ($s^{-1}$ $M^{-1}$) | 4.2 | 14.1 |

**Figure 2.1:** Schematic of *B. subtilis* biotin biosynthesis and the BioW reaction. A) The pathway begins with the conversion of pimelate to pimeloyl-CoA by an ATP-dependent reaction catalyzed by BioW, with the adenylate as intermediate. The pimelate thioester provides most the carbon atoms of the biotin, as indicated in red. The pimelate thioester reacts with L-alanine in a decarboxylative condensation catalyzed by BioF to form 8-amino-7-oxononanoate, the first intermediate in formation of the fused heterocyclic rings of biotin. Ring formation is completed by successive reactions catalyzed by BioA, BioD and BioB to generate biotin. B) The BioW reaction with pimelate as substrate with separation of substrates and products by reverse-phase HPLC. The absorbance peaks were identified using authentic standards of ATP, AMP and CoA. Below the chromatogram the steady state kinetic parameters for pimelate and CoA determined in sodium HEPES buffer (pH 7) with 50 nM BioW are given. The parameters were determined by fitting the data from three independent experiments to Lineweaver-Burk plots and are given as the mean ± SE. C) Thin layer chromatography of [α-$^{32}$P]ATP-labeled reactions in the absence and presence of CoA. A time course is shown for the CoA reactions. D) Identification of pimeloyl-CoA (calculated mass 910.18) by electrospray ionization mass spectroscopy. The sharp peak eluting after 31 min in B) was collected, dried under nitrogen and submitted for mass spectroscopy.

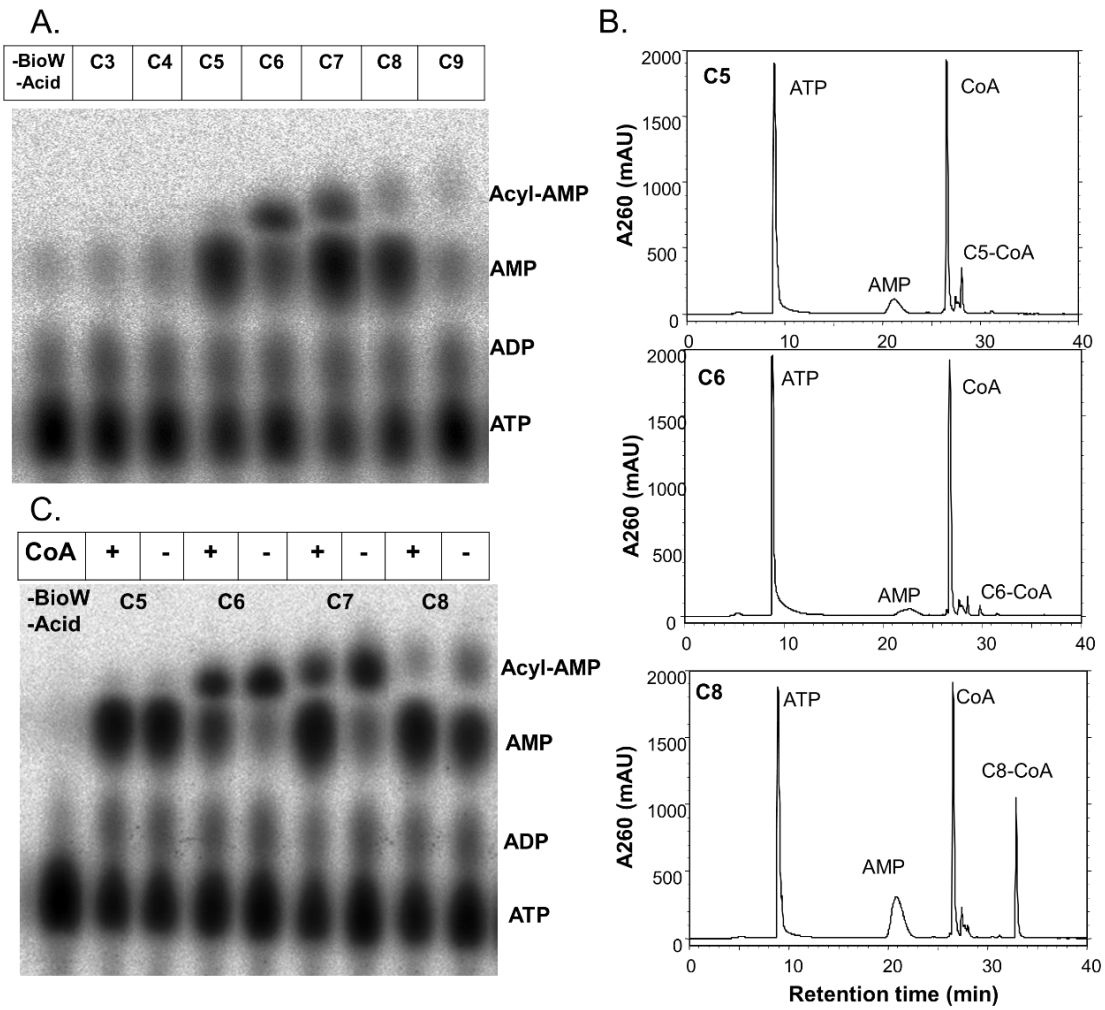



**Figure 2.2:** Dicarboxylate specificity of BioW. A) BioW reactions were performed with dicarboxylic acids from C3 to C9 to determine acyl chain substrate specificity. The reactions were carried out as described in Methods and Materials, and 1 µl of each reaction was spotted onto cellulose TLC plate. B) CoA thioester formation from the C5, C6 and C8 dicarboxylates after reaction with BioW at 37°C for 1 hr. Panels from top to bottom show HPLC analyses of BioW reactions with the C5, C6 and C8 dicarboxylates, respectively. C) Evidence of adenylate editing in absence of CoA. BioW reactions with the C5, C6, C7 and C8 dicarboxylates were carried out in presence or absence of CoA at 37°C for 1 h. Samples of 1 µl were spotted onto a TLC plate which was developed as described in Methods and Materials. Note that the migration rates of the acyl-AMP intermediates vary with the dicarboxylate moiety. C8-CoA was identified by mass spectroscopy (Figure 2-7) whereas the trace amounts of C5-CoA formed were insufficient for analysis.

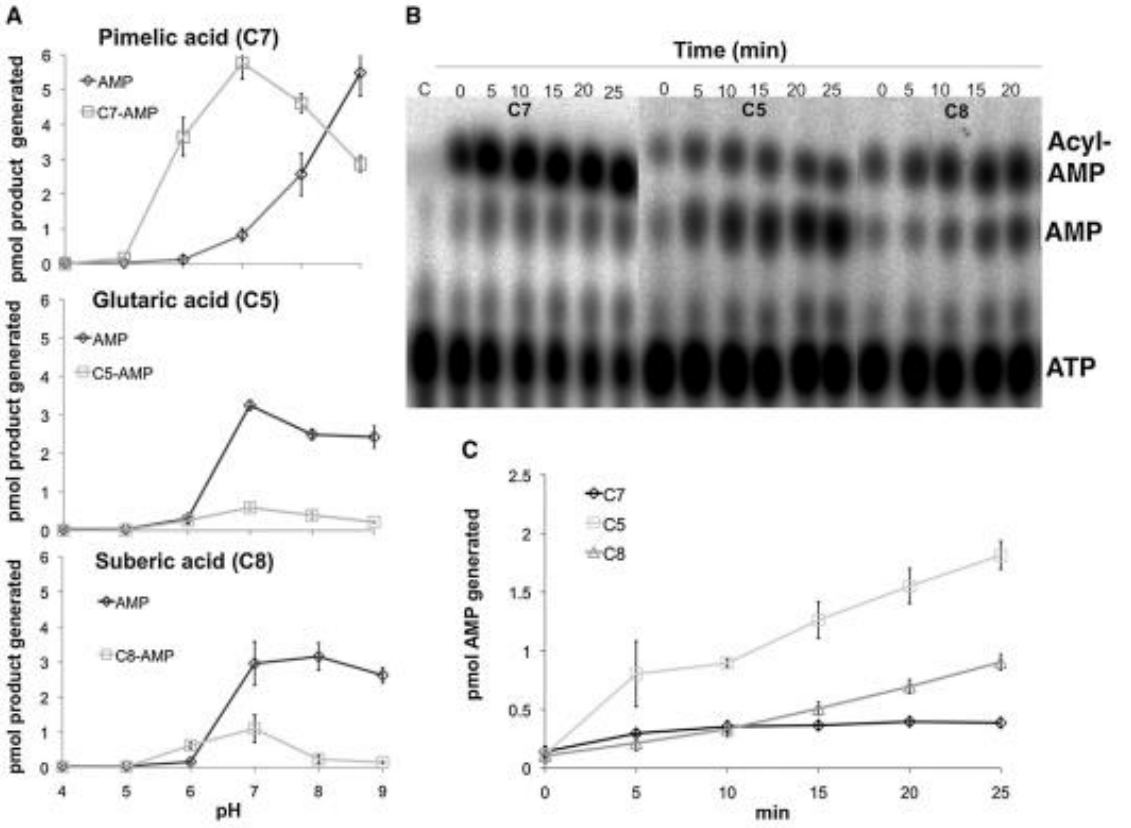



**Figure 2.3:** Cleavage of cognate and noncognate adenylates. A) The pH profiles of BioW activity with pimelate (top), glutarate (middle) and suberate (bottom) in absence of CoA. Buffers of sodium acetate (pH 4 and pH 5), sodium 2-(N-morpholino)ethanesulfonic acid (pH 6), sodium HEPES (pH 7) and Tris-HCl (pH 8 and pH 9) were used to determine the pH profile of the BioW reaction for 1 h at 37˚C. B) BioW catalyzed glutaryl- and suberyl-adenylate hydrolysis was assayed by determining the rate of AMP production with time. Reactions with pimelate (C7) were included for comparison. At different time points 1 µl samples of BioW reactions performed at pH 7.0 and 37˚C were spotted onto cellulose TLC plates followed by development and autoradiography. C) Quantitation of the data points in B) were used to show the rate of formation of AMP with increasing time with C7, C5 and C8 as substrates. The intensities of the spots were quantitated using the phosphor-imaging software of the Fujifilm FLA-3000 phosphorimager. The percent of photostimulated luminescence of AMP and acyl-AMP spots obtained from the reaction was compared to that of the ATP spot obtained from a control reaction lacking enzyme on each cellulose TLC plate to derive the amounts of products generated. All reactions were repeated three times, and the data are represented as SEs.

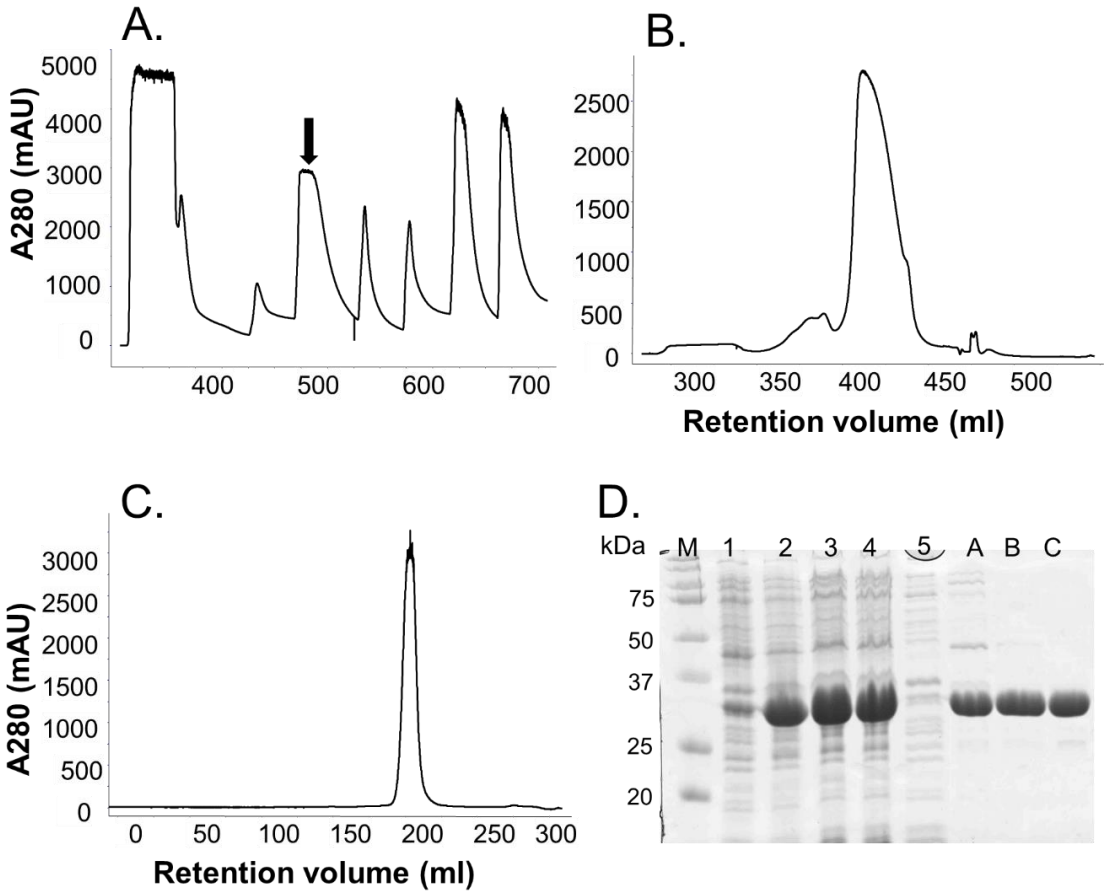



**Figure 2.4:** Three step purification of native *B. subtilis* BioW**.** A) Ion exchange chromatography using a HiTrap Q FF column. B) Hydrophobic interaction chromatography using HiTrap Phenyl HP column; C) Preparative size exclusion chromatography using HiLoad Superdex 200 column. D) 12% SDS-PAGE showing purity of BioW (29.7 kDa) in the three successive chromatographic steps denoted by A, B and C, respectively. *M,* molecular weight markers in kDa; lane 1, sample of uninduced culture; lane 2, sample of induced culture; lane 3, crude lysate; lane 4, supernatant; lane 5, flow-through of the ion exchange column.

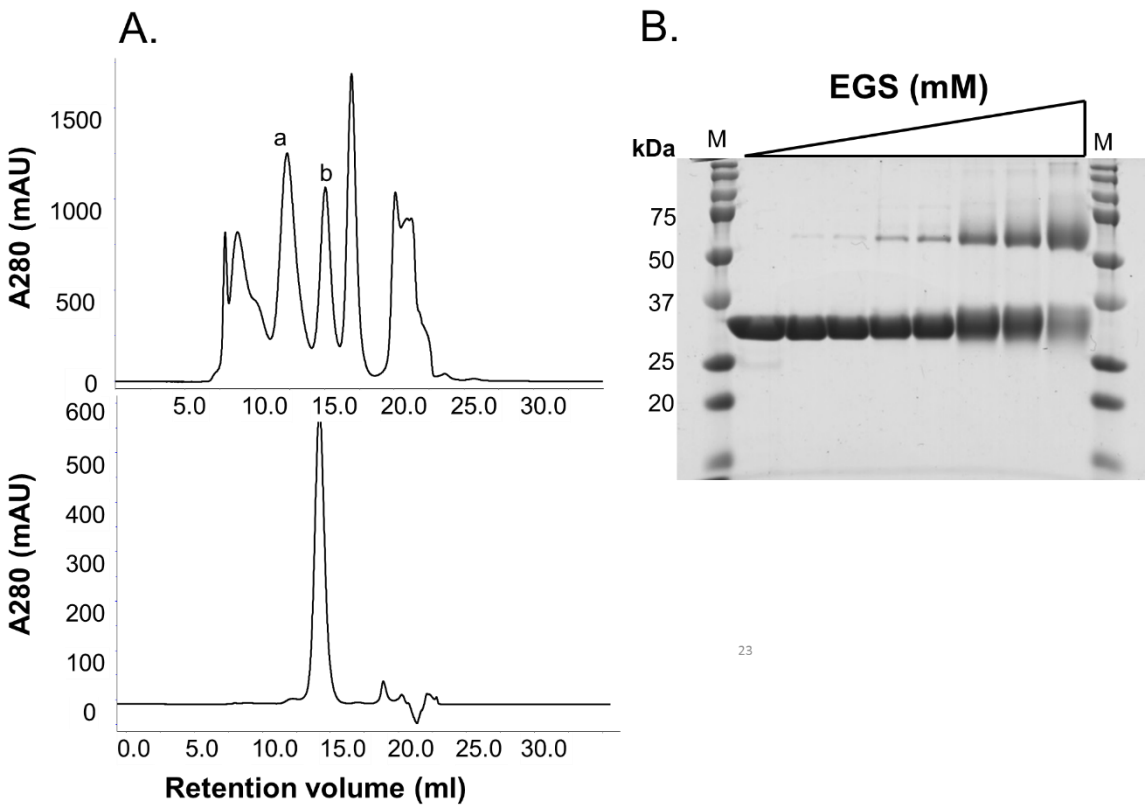



**Figure 2.5:** BioW solution structure**.** A) BioW (lower panel) eluted between bovine γ-globulin (a, 158 kDa) and chicken ovalbumin (b, 44 kDa) (upper panel) from an analytical Superdex 200 column, indicating that it is a dimeric protein. B) Chemical cross-linking of purified BioW (20 µM) with increasing concentrations (0, 0.005, 0.01, 0.05, 0.1, 0.5, 1 and 5 mM) of ethylene glycol bis-succinimidylsuccinate (EGS) shows a BioW dimer between 50 kDa and 75 kDa. The purified protein was incubated with EGS for 30 min at room temperature and the result was analyzed by 12% SDS-PAGE. M, protein molecular weight standards.

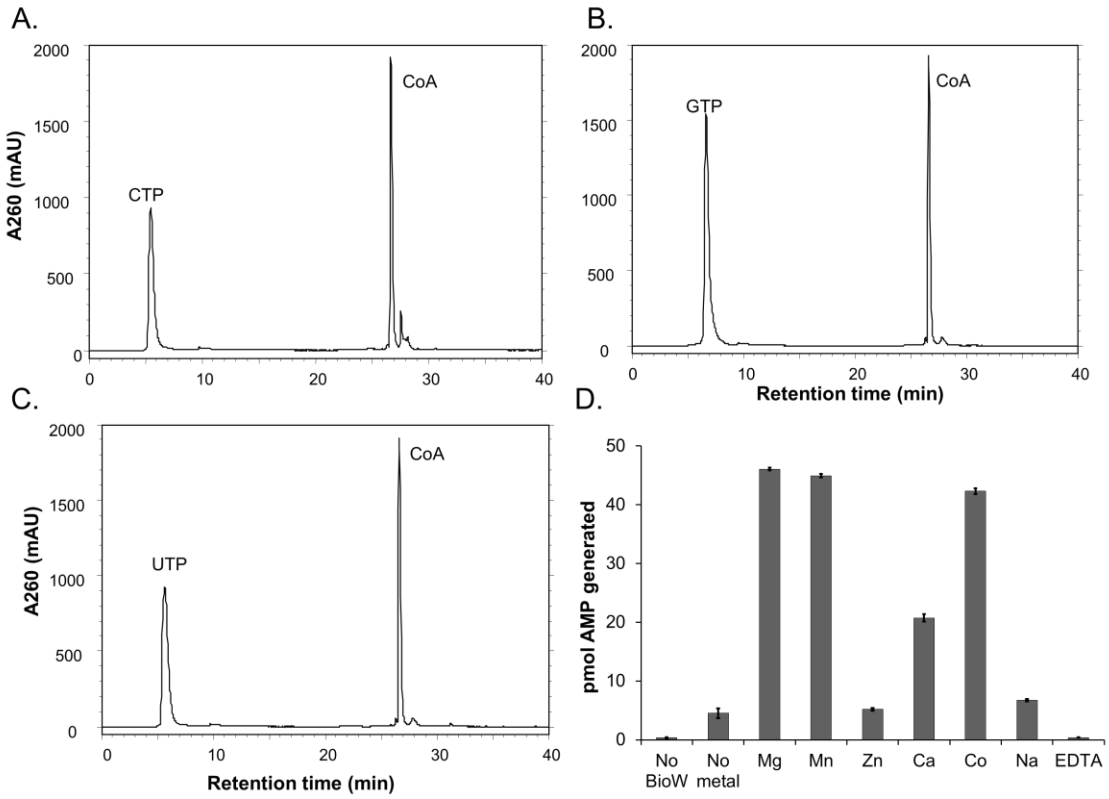



**Figure 2.6:** Nucleotide triphosphate and metal ion specificities of *B. subtilis* BioW. HPLC analysis of reaction of BioW with various nucleotide triphosphates. A) CTP, B) GTP and C) UTP. All panels show the CoA peak and a peak corresponding to the nucleotide triphosphate. The reaction and samples were handled as described in Methods and Materials. D), Metal ion dependence in reactions containing either a metal ion at 5 mM or 5 mM EDTA in a buffered solution of pH 7. All metal ions were added as the chlorides. The error bars represent standard deviations.

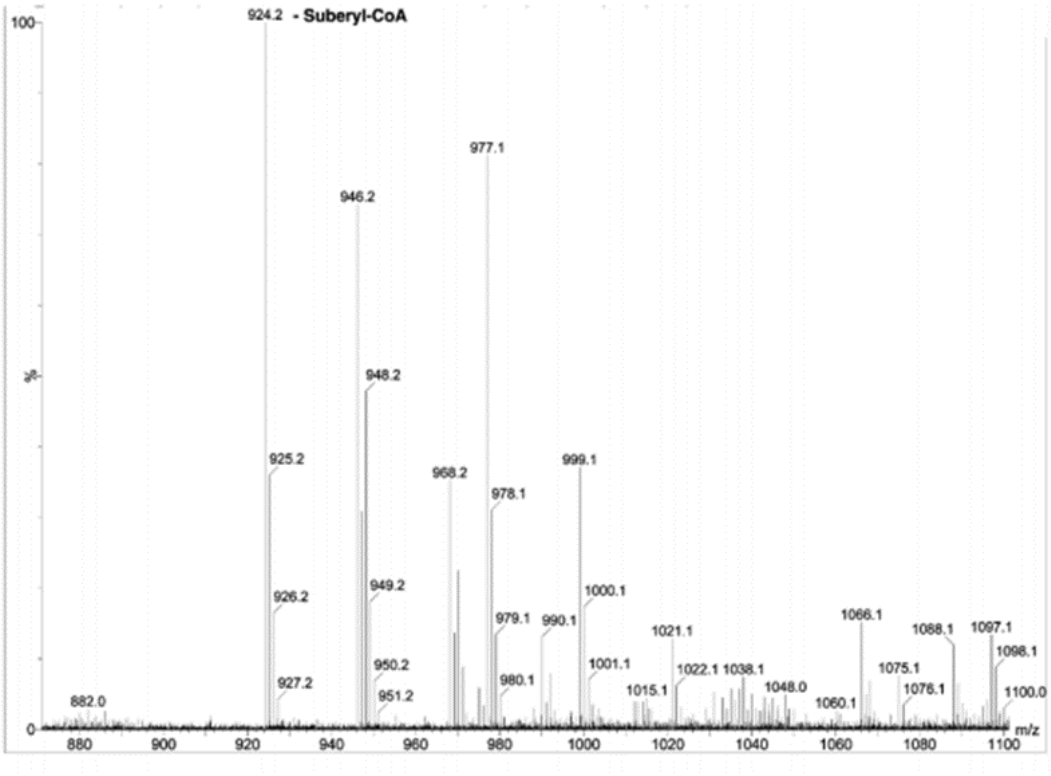



**Figure 2.7:** Identification of suberyl-CoA (calculated mass 924.2) by electrospray ionization mass spectroscopy. The analysis was performed as described in Methods and Materials on the HPLC peak of Figure 2-2B (C8-CoA). The peaks at higher mass are probably metal ion adducts of the acidic CoA moiety.

# CHAPTER 3

# PIMELIC ACID, THE FIRST PRECURSOR OF THE *BACILLUS SUBTILIS* BIOTIN SYNTHESIS PATHWAY, EXISTS AS THE FREE ACID AND IS ASSEMBLED BY FATTY ACID SYNTHESIS

## INTRODUCTION

The initial stage of biotin synthesis pathway varies widely depending on the enzymes responsible for synthesis of a thioester of pimelic acid, a seven carbon α, ω-dicarboxylic acid (IUPAC name, heptanedioic acid). The role of this unusual molecule as a biotin precursor was discovered many years ago as a compound that bypassed the biotin requirement for growth of *Corynebacterium diphtheriae* (91, 90, 89). Later work showed that pimelic acid supplementation bypassed the biotin requirements of a subset of naturally biotin-requiring bacteria and fungi. Conversion of $^{14}$C-pimelic acid into biotin and synthesis of biotin vitamers from pimelic acid in cell-free extracts of various bacteria indicated that pimelate was a direct precursor of biotin (58) and was probably incorporated intact eisenberg incorp (31). Seminal *in vitro* studies by Eisenberg and coworkers (30, 29) demonstrated that *E. coli* extracts could condense pimeloyl-CoA with alanine to give the first intermediate in biotin ring formation and that BioF catalyzed this reaction. However, the source of pimelate and the identity of its thioester derivative in *E. coli* remained unknown (the high pimeloyl-CoA concentrations required suggested it may not be the physiological thioester).

Recent work from this laboratory showed that synthesis of the acyl carrier protein (ACP) thioester of pimelate (pimeloyl-ACP) proceeds by a modified fatty acid synthesis pathway(1, 76, 75). Pimeloyl-ACP then reacts with alanine in the BioF reaction to give 8-amino-7-oxononanoate. Although many bacteria use the BioC-BioH pathway (or variations in which other esterases replace BioH), there are biotin prototrophic bacteria that lack BioC encoding genes and a known esterase able to cleave pimeloyl-ACP methyl ester. In this regard the most puzzling of these are the Bacilli. The *bio* operons of *B. subtilis* and closely related bacilli (e.g. *B. licheniformis, B. amyloliquefaciens,* etc.) lack

*bioC* and *bioH* and instead have *bioW* and *bioI*. In contrast *B. cereus* and closely related bacteria (e.g., *B. anthracis*, *B. thuringiensis*, etc.) encode BioC and BioH proteins within their biotin operons that are quite similar and can functionally replace the *E. coli* proteins.

The *B. subtilis bioW* and *bioI* genes are, respectively, the first and last genes of the biotin operon (10) (Figure 3-1). The intervening genes, *bioAFDB,* are readily identified due to the strong conservation of their protein sequences across diverse bacteria and fungi. *B. subtilis* BioW has long been known to be a pimeloyl-CoA synthetase (10, 107) and this atypical acyl-CoA synthetase has been studied in detail in recent enzyme mechanism and x-ray crystal analyses (33, 82, 127). In Chapter 2, I have shown the enzymatic function of BioW as a pimeloyl-CoA synthetase with a unique proofreading activity. BioI is characterized as a cytochrome P450 due to its sequence homology and characteristic P450 spectral properties (43, 119). BioI has been shown to cleave carbon-carbon bonds of free fatty acids to mixtures of dicarboxylic acids that include pimelate(119). BioI has also been reported to cleave the carbon-carbon bonds of acyl-ACPs that co-purified with preparations of BioI expressed in *E. coli* (119) to give pimeloyl-ACP, albeit at sub-stoichiometric levels. The prior work plus the elegant crystal structures argued strongly that BioI provided the pimelate moiety required for *B. subtilis* biotin synthesis. Hence, *B. subtilis* seemed to have redundant systems in the first stage of the pathway. However two considerations weakened this argument. First early genetic studies concluded that BioI function was not strictly required for biotin synthesis (10) and bacteria exist that encode BioW but lack genes encoding a recognizable BioI. I found that disruption of *bioI* fails to result in biotin auxotrophy whereas disruption of *bioW* results in a growth requirement for biotin. The auxotrophy of *bioW* mutant strains plus the BioW pimeloyl-CoA synthetase activity argued that free pimelic acid must be a biotin precursor. I report isotopic labelling studies that demonstrate this is the case and that the pimelate moiety is synthesized by the head-to-tail condensations of acetate units characteristic of fatty acid synthesis.

# MATERIALS AND METHODS

## Materials and strains

Sodium acetate (1-$^{13}$C and 2-$^{13}$C, each 99% enriched) was purchased from Cambridge Isotope Laboratories, Inc. Cerulenin, platencin, platensimycin and 8-amino-7-oxononanoate were purchased from Cayman Chemical. All other chemicals were purchased from Fisher Scientific or Sigma Aldrich. The bacterial strains are given in Table 3-1. The *B. subtilis* strains are all derivatives of strain 168 whereas the *E. coli* strains are derivatives of the wild type K-12 strains MC1061 or MG1655. The *B. subtilis* biotin overproducing strain BI274 and BI304 (derivatives of strain PY79 with constitutive expression of the *bio* operon) (11) were from the American Type Culture Collection (accession 55575 and 55573 respectively). The cer-20 derivative of *B. subtilis* 168 that is partially resistant to cerulenin (115) was from the *Bacillus* Genetics Stock Center (accession 1A577).

## General growth media

*E. coli* and *B. subtilis* strains were grown in LB media. The minimal medium for *E. coli* contained M9 salts, 0.4% glucose, 1 µg/ml thiamine and 0.1% Casamino acids. General defined media for *B. subtilis* was Spizizen salts consisting of $(NH_4)_2SO_4$; 0.2 %, $K_2HPO_4$; 1.4%, $KH_2PO_4$; 0.6%, sodium citrate · $2H_2O$; 0.1%, $MgSO_4 \cdot 7H_2O$; 0.02% (118), plus trace elements [$MgCl_2 \cdot 6H_2O$, $CaCl_2$, $FeCl_2 \cdot 6H_2O$, $MnCl_2 \cdot 4H_2O$, $ZnCl_2$, $CuCl_2 \cdot 2H_2O$, $CoCl_2.6H_2O$ $NaMoO_4.2H_2O$], 0.5% glucose, 0.04% potassium glutamate and 1 mM $MgSO_4 \cdot 7H_2O$. Supplements required for growth included 0.01% tryptophan, 1 mM sodium acetate, 0.1 mM each of branched chain fatty acid (BCFA) precursors (isobutyric acid, isovaleric acid and 2-methylbutyric acid). Genetic complementation experiments requiring expression of Pspac and $P_{BAD}$ promoters were performed in minimal media with 0.5% glycerol as the carbon source. Either isopropyl β-D-1-thiogalactopyranoside (IPTG) (1 mM) or 0.2% arabinose was used for induction of the respective promoters. Antibiotics were used in the following concentrations (in µg/ml): chloramphenicol (5), erythromycin sulfate (1), lincomycin hydrochloride (25), spectinomycin sulfate (100), sodium ampicillin (100) and kanamycin sulfate (50).

Construction of *bioW* and *bioI* deletion mutant strains

A 432 bp internal *bioW* fragment was amplified by PCR using primers: 5` GAG ATC GAA TTC CCA TTC AGC CAT TGC 3` and 5` ATA ATG CGG CCG CTC ATT GCT GTA ATA CG 3`. The PCR product was digested with EcoRI and NotI and inserted into pMUTIN4 (123). The resultant recombinant plasmid was used to transform *B. subtilis* 168 to allow single crossover recombination of the nonreplicating plasmid into the chromosome. The transformants were selected on plates containing erythromycin and lincomycin. To disrupt *bioI* fragments of 570 bp and 627 bp corresponding to the 5` and 3` region of *bioI,* respectively, were amplified with primers 5` ATC AGG GAT CCA AAT GAA GGC TAG TTT AAG 3`, 5` CAC TCC GAA TTC TGC TCC CTA TCT TCC 3` (5` region) and 5` CGG AC GTC GAC CAA CGG TCA ATC TCA TC 3`, 5` AAG CAC GGG CCC TTC ATA GTC TGA AAT AAG C 3`(3` region) and inserted into the BamHI-EcoRI (5`) and SalI-ApaI (3`) sites that bracket the kanamycin cassette of pDG780 (44). The recombinant plasmid encodes a *bioI* in which 75 internal codons were replaced with the kanamycin cassette. This plasmid was used to transform *B. subtilis* 168 and the recombinants resulting from a double-crossover recombination were selected on kanamycin-containing LB plates. Transformation of *B. subtilis* was carried out by the procedure of Dubnau and Davidoff-Abelson (27).

For ectopic expression of *bioW,* the coding region was inserted into the SalI-SphI sites of pDR111. The recombinant plasmid was introduced into the *bioW* mutant strain by a double crossover recombination at the *amyE* locus. The gene disruption constructs were verified by colony PCR, restriction digestion and sequencing of the PCR products. The *amyE* phenotype was assayed on colonies grown for 16 h on LB starch plates by flooding the plates with 1% $I_2$-KI solution (116).

Growth of strain MM194 and purification of biotin from culture supernatants

Strain MM194 was constructed by transformation of strain BI274 with the genomic DNA of NM57 (JH642 Δ*lipM*::Km) (84) with selection for *ΔlipM* mutants on kanamycin plates. The mutants were verified by the growth phenotype on minimal media and colony PCR. Strain MM194 was cultured overnight in LB and subcultured at the

final $OD_{600}$ of 0.02 in minimal medium I which contained 0.05% yeast extract and 0.02% Bacto-Tryptone and allowed to grow at 37°C for 6-9 h until an $OD_{600}$ ~2 was reached. Finally, the MMI culture was transferred to 1 L of minimal medium II which lacked yeast extract and tryptone but contained sodium acetate and BCFA at a final $OD_{600}$ of 0.02 and allowed to grow overnight at 37°C for 15-18 h.

Cells were removed from 10-20 L cultures by centrifugation and the supernatant was filtered through Millipore 0.22 µm membranes. The supernatant was treated with activated charcoal and the charcoal was collected and extracted with an ethanol-ammonium hydroxide mixture as described by Ogata (96). This extract was concentrated to a volume of 2 mL using Buchi Rotavapor R-210. N-hydroxysuccinimide-activated agarose (Pierce Thermo-Fisher Scientific) was used to covalently immobilize the F43A mutant derivative of shwanavidin, a dimeric avidin from *Shewanella denitrificans* (86) following the manufacturer protocol. Shwanavidin (20 mg) was coupled to 150 mg of agarose in coupling buffer (100 mM $Na_2HPO_4$, 150 mM NaCl at pH 7.2) by mixing at 4 °C overnight. The shwanavidin-coupled agarose was washed twice in coupling buffer and any remaining N-hydroxysuccinimide groups were quenched with 1M Tris-HCL (pH 7.4). Multiple loadings of 200-400 nmol biotin in the charcoal extract samples were allowed to bind to the shwanavidin-agarose complex by mixing at 4°C overnight. The agarose was then washed three times with coupling buffer. Most of the colored impurities were found in the flow-through and wash fractions with some loss of biotin. The bound biotin was then eluted with heated water (70°C) (52). Biotin was repeatedly determined by bioassay throughout the process of purification. The biotin samples obtained (~1 mg from 10 L cultures) were analyzed by Agilent 600 MHz NMR at the Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign. $D_2O$ was used as solvent (pH ~13) and number of scans was 16,000nt. The biotin carbon atoms were assigned based on the $^{13}C$ NMR shifts independently determined by Sanyal et al. (113) and Ifuku et al.(55).

Shwanavidin purification

A synthetic gene encoding the F43A mutant derivative of *Shewanella denitrificans* shwanavidin was inserted into the NdeI-HindIII sites of vector pET28b to encode a protein with a N-terminal hexahistidine tag. The recombinant plasmid was transformed into expression strain BL21 (DE3) Tuner (Novagen) with selection for kanamycin resistance to obtain strain MM200. Strain MM200 was grown overnight in LB medium and subcultured into 1 L of LB to allow growth up to $OD_{600}$ of 0.8. The culture was induced with 1 mM IPTG for 3 h at 37°C. The cells were harvested after centrifugation and washed with 10 mM Tris-HCl, 100 mM NaCl and 1 mM EDTA (pH 7.5). Cell pellets were suspended in 50 mM Tris-HCl, 200 mM NaCl, 1% Triton X-100, 8% sucrose and 1mM phenylmethylsulfonyl fluoride at pH 8. Cells were lysed by passage through a French press and the lysate was centrifuged at 18,000 rpm for 30 min. The insoluble inclusion bodies were washed in the same buffer without Triton X-100 and centrifuged again to remove the buffer. The inclusion bodies were then dissolved in 6 M guanidine hydrochloride and 50 mM Tris-HCl (pH 7.5) to a concentration of 10 mg /mL and stirred slowly at 4°C for 4-6 h to allow equilibration. The solubilized protein was added dropwise into folding buffer (1/4 -1/2 of induced culture volume) of 50 mM Tris-HCl, 150 mM NaCl, 5 mM EDTA and 0.1 mM phenylmethylsulfonyl fluoride at pH 7.5 at 4°C while stirring and allowed to equilibrate overnight. The solution was centrifuged to remove insoluble debris and dialyzed against 50 mM $NaH_2PO_4$, 500 mM NaCl and 10% glycerol (pH 8). The partially pure protein was further purified using an ÄKTA protein purification system and a 5 mL HisTrap HP column. The pure protein was eluted with 150 mM imidazole in 50 mM $NaH_2PO_4$, 500 mM NaCl and 10% glycerol (pH 8). The protein was concentrated using 3K Amicon concentrator (EMD Millipore) and flash frozen at -80 °C in 50 mM $NaH_2PO_4$, 500 mM NaCl and 10% glycerol (pH 8).

Mass spectral analyses

Purified biotin samples were analyzed with Agilent LC/MS (Agilent Technologies, Santa Clara, CA). The LC separation was performed on an Agilent Eclipse XDB-C18 (4.6 x 50mm, 5µm) with mobile phase A (0.1% formic acid in water) and mobile phase B (0.1% formic acid in acetontrile). The flow rate was 0.4 mL/min. The

linear gradient was as follows: 0-2 min, 100% A; 6-17.5 min, 0% A. For high resolution mass spectra, the samples were analyzed by using the Q-Exactive MS system (Thermo, Bremen, Germany) of the Metabolomics Laboratory of the Carver Biotechnology Center, University of Illinois at Urbana-Champaign.

Bioassay of biotin production

*E. coli* strain NRD25 (15), which lacks the entire *bio* operon and strain ER90 (Δ*bioF bioC bioD)*, which retains a functional *bioB* (14) were grown overnight in defined media containing 1 nM biotin. The cells were harvested by centrifugation, washed three times in M9 salts and allowed to starve in 100 mL of defined media containing 0.5 units of avidin for 5 h at 37°C. The biotin-starved cells were washed three times in 1mL M9 salts and mixed with melted 1.5% agar containing defined media and 0.01% 2,3,5-triphenyltetrazolium chloride, a redox indicator that develops a red color upon cellular respiration (17, 75). About 5 ml of medium was poured into plates quarter-sectored by plastic walls and sterile 6 mm disks (BD BBL blank test discs) was placed on top of the solidified agar. After spotting 10 µl of sample onto the disks the plates were incubated at 30°C overnight. The amounts of biotin in the samples were quantitated by measuring the diameter of the red zone produced by known amounts of commercial biotin used as controls for every batch of minimal plates.

Resting cell assays

Strain BI274 cer-20 was constructed by transforming strain BI274 with the genomic DNA of the original cerulenin resistant strain and selection on LB plates containing 10 µg/ml cerulenin. The FabF mutation conferring increased cerulenin resistance (I108F) (115) was verified by sequencing. Strain BI274 and BI274 cer-20 were grown in LB media for 7-8 h at 37°C. The cells were subcultured to 400 mL of defined minimal media to a final $OD_{600}$ of approximately 0.02 and incubated at 37°C for 16 h. The cells were harvested by centrifugation at 8000 rpm for 10 min and washed twice with Spizizen salts. Cells were transferred to 25 mL nitrogen-deficient minimal medium (Spizizen salts lacking ammonium sulfate) to a final $OD_{600}$ of approximately 3 to 5. Aliquots were taken at several time points after adding antibiotics and the amount of

biotin released into the supernatant was determined by bioassay with *E. coli* strain ER90 (*ΔbioF bioC bioD)* (14). The amount of biotin produced was quantitated by comparison to known amounts of control biotin assayed on the same batch of plates.

## RESULTS

### Expression of either *B. subtilis* BioW or BioI bypasses the biotin auxotrophy of *an E. coli ΔbioC* Δ*bioH* strain

The *B. subtilis bioW* and *bioI* coding sequences were inserted into the *araBAD* promoter vector pBAD322 (18) and the plasmid was introduced into strain STL25, a Δ*bioC* Δ*bioH* derivative of strain MG1655. Complementation of the *E. coli ΔbioC* Δ*bioH* strain was tested in biotin-free minimal medium supplemented with pimelic acid. The biotin auxotrophy of the mutant was successfully rescued by the induction of *bioW* gene expression with arabinose but only in the presence of pimelic acid (Figure 3-2A) as previously reported (10, 75). In contrast, basal level expression of *bioI* without arabinose induction was able to bypass the *ΔbioC* Δ*bioH* mutations even in the absence of pimelic acid, consistent with results obtained by Bower et al. (10) (Figure 3-2B).

### *bioW* is essential for biotin synthesis whereas *bioI* is dispensable

*B. subtilis* seems to contain two genes that serve the same purpose, generation of the pimelate thioester intermediate. In order to determine the importance of *bioW* and *bioI*, I constructed chromosomal gene disruptions. Vector pMUTIN4 was used to inactivate the chromosomal *bioW* gene by single-crossover homologous recombination and has an IPTG-inducible promoter (Pspac) positioned to drive expression of the downstream genes (123). Disruption of *bioW* caused biotin auxotrophy in biotin-free media (Figure 3-3A). Growth was rescued by expression of *bioW* from an IPTG-inducible promoter at an ectopic chromosomal site (the *amyE* locus) (Figure 3-3A) demonstrating that the biotin auxotrophy is due to *bioW* inactivation and is not a polarity effect. Also, supplementation with dethiobiotin or 8-amino-7-oxononanoate rescued growth of the Δ*bioW* strain indicating that induction of the Pspac promoter prevented polar effects on expression of the downstream genes.

In contrast, disruption of *bioI* failed to cause biotin auxotrophy. The strain grew as well in biotin-free minimal media as the wild type strain (Figure 3-3A). It should be noted that the internal fragment deleted from *bioI* encodes several residues that position the acyl chain for attack by heme moiety (19). In the first studies of the *B. subtilis bio* operon insertion of an antibiotic resistance cassette into a site early in the BioI coding sequence resulted in strains that grew (albeit very slowly) in biotin-free minimal media (10). Moreover, a Δ*bioW* derivative of strain BI304 (11, 104) encoding a modified *bio* operon in which the terminator upstream of *bioI* was replaced with a constitutive strong promoter resulting in overexpression of BioI (104) failed to grow in biotin-free media. However, upon expression of a *bioW* gene integrated at the ectopic *amyE* locus this strain grew well (Figure 3-4).

To provide a further test of the nonessential nature of BioI in *B. subtilis* biotin synthesis I tested the ability of the wild type strain 168 to grow anaerobically with nitrate as electron acceptor (93). The cytochrome P450 reaction requires molecular oxygen (81) and thus if BioI played an essential role in biotin synthesis growth should be blocked under these conditions. Two different anaerobic systems were tested, the BD Gaspak EZ Anaerobe Pouch System and a Coy anaerobic chamber containing 10% $H_2$ - 90% $N_2$. Growth proceeded in both systems regardless of the presence of biotin demonstrating that oxygen is not required for biotin synthesis (Figure 3-3B). Nitrate was required for growth confirming that the culture conditions were anaerobic.

Free pimelic acid is a *bona fide* precursor of biotin synthesis in *B. subtilis*

The origin of pimelic acid in bacteria lacking the *E. coli* pathway remains a mystery. I carried out $^{13}$C-NMR experiments to test for the presence of free pimelic acid *in vivo* in *B. subtilis* and to begin to deduce a pimelate synthesis pathway. These experiments were analogous to previous studies in *E. coli* by Sanyal, Ifuku and coworkers (55, 113). In both studies, *E. coli* was grown in the presence of [1-$^{13}$C]acetate or [2-$^{13}$C]acetate and the labelling patterns of biotin extracted from the media was analyzed by $^{13}$C nuclear magnetic resonance (NMR). The authors ruled out free pimelic

acid as a biotin precursor in *E. coli* because the labelling patterns of biotin (and dethiobiotin) demonstrated that one of the pimelate carboxyl groups was fixed during synthesis possibly by a thioester bond. (If free pimelate was an intermediate, the biotin carbon atoms derived from the pimelate carboxyl groups would show the same labeling patterns due to the symmetry of pimelate.) Given that BioW is essential for biotin synthesis and activates free pimelate I expected a different result. Free pimelic acid is a rotationally symmetric molecule and thus the carbon atoms derived from the carboxyl groups would have the same labelling pattern.

In this work I used *B. subtilis* strain BI274 which has an engineered *bio* operon driven by a phage SP01 promoter resulting in overproduction of biotin up to 1-2 mg/L of growth medium (11). *B. subtilis* cannot grow on acetate as sole carbon source because of the lack of a glyoxylate cycle and thus had to be grown on a more complex carbon source. In order to prevent dilution of the supplemented $^{13}$C-acetate by acetate produced from the glucose carbon source, I used a Δ*lipM* derivative of strain BI274 called strain MM194. The loss of LipM blocks lipoic acid synthesis which inactivates the essential lipoic acid-requiring enzymes, pyruvate dehydrogenase and branched-chain dehydrogenase (84). However, loss of these enzymes can be bypassed by supplementation with acetate and branched chain fatty acid precursors which endogenous CoA ligases convert to their CoA esters (84).

Strain MM194 was grown with either [1-$^{13}$C]acetate or [2-$^{13}$C]acetate as the sole acetate source to allow maximum enrichment of the $^{13}$C label into biotin (the branched chain precursors were also added) (Figure 3-5A). I then attempted to use the methods developed for *E. coli* to extract and purify biotin from the culture medium in sufficient amounts and purity to perform $^{13}$C nuclear magnetic resonance (NMR) analyses. However, although the recovery of biotin was good (Figure 3-5B), the resulting samples were much too impure for NMR. Several additional purification steps were unsuccessful in ridding the samples of intensely colored contaminants. Finally I found that a mutant (F43A) derivative of shwanavidin, a dimeric avidin from *Shewanella denitrificans* (86) provided the needed purification. In native shwanavidin phenylalanine-43 acts to shield

bound biotin from solvent and its substitution with alanine results in a precipitous drop in the affinity of the protein for biotin to $10^{-8}$ M (86). Expression in *E. coli* of a synthetic gene encoding the mutant shwanavidin gave abundant protein that was refolded immobilized on agarose. The impure biotin samples were bound to the shwanavidin F43A column, washed with three column volumes of 100 mM $Na_2HPO_4$ containing 150 mM NaCl (pH 7.2) to remove the colored contaminants and biotin was eluted with water heated to 70°C. The elution conditions were those used to release biotinylated nucleic acids from streptavidin (52). Shwanavidin like other avidins is very heat tolerant and thus the column could be reused multiple times.

As shown in Figure 3-6B, carbons 3, 10, 12 and 14 were enriched when [1-$^{13}$C]acetate was fed whereas when [2-$^{13}$C]acetate was the acetate source carbons 7, 11 and 13 were enriched (Figure 3-6C). Simultaneous $^{13}$C enrichment from [1-$^{13}$C]acetate of the biotin carbon atoms derived from the pimelate carboxyl groups, carbons 3 and 14, provides strong evidence for the existence of free pimelate, a symmetrical molecule. In addition, the alternating pattern of enrichment of carbons from [1-$^{13}$C]acetate and [2-$^{13}$C]acetate indicates that pimelate synthesis involves the head-to-tail chain elongations characteristic of fatty acid synthesis. The enrichments of the $^{13}$C-biotin samples used for NMR analysis were verified by high resolution electrospray ionization mass spectroscopy (Figure 3-7). The biotin samples labeled with [1-$^{13}$C]acetate and [2-$^{13}$C]acetate contained four and three $^{13}$C atoms, respectively, demonstrating that all of the acetate was derived from the supplements. The lack of dilution of the $^{13}$C-acetate supplements by $^{12}$C-acetate produced from the glucose carbon source indicates that the Δ*lipM* mutation completely inactivated the pyruvate dehydrogenase of strain MM194.

Synthesis of pimelate depends on fatty acid synthesis in *B. subtilis*

In order to further explore the pimelate link between fatty acid and biotin synthesis, I carried out resting cell assays with fatty acid enzyme inhibitors and monitored the effects on biotin synthesis. I tested the effects of cerulenin (115), platensimycin and platencin, inhibitors of fatty acid elongation enzymes, FabF/FabB and FabH/FabF, respectively (131). Biotin production was determined by bioassay. Resting

cell assays (cells starved for nitrogen) were used to avoid the difficulties of comparing inhibitor-free growing cultures to inhibited cultures that were unable to grow due to lack of lipid synthesis.

Cerulenin at 10 µg/ml was found to inhibit the production of biotin by resting cells of strain BI274 by more than 70% (Figure 3-8) whereas 40 µg/ml cerulenin inhibited biotin production by more than 80% (Figure 3-8) indicating that the generic fatty acid elongation enzyme FabF, is required for biotin synthesis. Biotin synthesis was restored to 50% (at 10 µg/ml cerulenin) in the FabF [I108F] cerulenin-resistant derivative of strain BI274 (Figure 3-8) which is relatively insensitive to low concentrations of cerulenin (115). In addition, pimelic acid supplementation of cerulenin-treated resting cell cultures resulted in restoration of biotin synthesis to levels comparable to that of the untreated cultures whereas supplementation with glutaric acid had no effect (Figure 3-8). Therefore, the effects of cerulenin on biotin synthesis were due to blockage of pimelate synthesis.

I also tested two other fatty acid synthesis inhibitors that to my knowledge had not been tested on *B. subtilis.* These were a dual FabH/FabF inhibitor, platencin (126), which inhibited biotin production by 19% when added at 1.9 µg/ml and platensimycin a FabF specific inhibitor (126), which inhibited biotin production by 15% at the same concentration (Figure 3-9). Note that the platencin and platensimycin experiments were limited by the expense of the antibiotics coupled with the relative insensitivity of *B. subtilis* to these molecules.

## DISCUSSION

The initial stage of *B. subtilis* biotin synthesis is markedly different from that of *E. coli* in that both the enzymes and thioester moieties differ (CoA in *B. subtilis,* ACP in *E. coli*). Although the presence of *bioI* and *bioW* in *B. subtilis* indicated a redundancy in the generation of the pimelate thioester intermediate, my data shows that *bioW* is the only *bio* gene essential for generation of the pimeloyl-CoA substrate of BioF. Bower *et al.*

reported that inactivation of *bioW* resulted in a biotin-requiring phenotype (10). However, growth was only partially restored by addition of dethiobiotin or 7, 8-diaminononanoate indicating polar effects on the downstream genes. In the *bioW* disruption construct I included a promoter oriented to transcribe the downstream genes. This was successful in that supplementation with late intermediates in the pathway completely restored growth. Moreover, complementation of the Δ*bioW* mutation by expression of wild type *bioW* from an ectopic site further demonstrated the absence of polar effects on the downstream genes.

My $^{13}$C-NMR data provide the first definitive evidence for the presence of free pimelic acid in *B. subtilis* cells. Following the rationale of Sanyal *et al.* and Ifuku *et al.* (55, 113) that eliminated free pimelic acid as a biotin precursor in *E. coli*, I conclude that the opposite situation is the case for *B. subtilis.* In the *E. coli* studies, free pimelic acid was ruled out as the precursor because upon incorporation into biotin the carbon atoms derived from the carboxyl groups of pimelate were not labelled symmetrically and thus are metabolically distinct. In contrast, my data showed that feeding experiments with [1-$^{13}$C]acetate enriched both of the biotin carbon atoms derived from the pimelate carboxyl groups indicating that free pimelic acid, a symmetrical molecule, is the precursor. In addition, my data also indicate fatty acid synthesis to be responsible for pimelate synthesis. This conclusion is based on the alternate pattern of assimilation of labelled carbons from [1-$^{13}$C]acetate and [2-$^{13}$C]acetate which indicate a head-to-tail configuration identical to the fatty acid chain elongation process. Resting cell assays with FabF inhibitors decreased biotin production indicating that fatty acid synthesis is involved in pimelate and hence, biotin synthesis. The effect of cerulenin on biotin production was more pronounced than with the other two inhibitors tested (note that *B. subtilis* is naturally resistant to two other fatty acid synthesis inhibitors, thiolactomycin and triclosan). I expect that the efficiency of cerulenin in inhibiting biotin production is probably due to its covalent interaction with FabF whereas platencin and platensimycin bind non-covalently (63). Moreover, platencin and platensimycin are poor inhibitors of *B. subtilis* growth. The concentrations reported to inhibit *Staphylococcus aureus* were ineffective with *B. subtilis*. It should be noted that even inhibitors that almost completely

block fatty acid synthesis (e.g., cerulenin) might be expected to have less effect on biotin synthesis due to the fact that any residual fatty acid synthesis capacity could suffice for biotin synthesis given that fatty acids are made in great excess (*ca.* $10^4$) over biotin. Note that I made numerous attempts to develop an *in vitro B. subtilis* cell extract system that would synthesize dethiobiotin starting from malonyl-CoA. A successful system would have allowed us to test intermediates and to overcome any cellular impermeability to inhibitors as was done using *E. coli* extracts (75). However, this approach was unsuccessful, at best I saw only traces of dethiobiotin synthesis. Indeed, prior workers reported that *B. subtilis* extracts were almost totally deficient in synthesis of fatty acids that lacked terminal branching (13) which may have precluded synthesis of the straight chain pimelate molecule,.

The ability of *B. subtilis* BioI to support robust growth of *E. coli* Δ*bioC* Δ*bioH* strains in the absence of biotin indicated that the gene encodes a functional enzyme that almost certainly generates pimeloyl-ACP, a known intermediate in the *E. coli* pathway that readily explains the Δ*bioC* Δ*bioH* bypass. In *B. subtilis* the *bioI* transcript is 8-fold less abundant than the upstream *bio* operon transcript due to transcription termination (104). However, deletion of the terminator had only a very modest effect (<2-fold) on biotin production (104) and did not allow biotin synthesis in a Δ*bioW* strain. Hence, BioI does not play a detectable role in *B. subtilis* biotin synthesis even when overproduced and thus presents an enigma. Bower and coworkers reported that a *ΔbioI* mutation caused a bradytrophic phenotype in biotin-free media (10). However, I found that disrupting *bioI* failed to cause any growth defect. The ability of *B. subtilis* to grow anaerobically in biotin-free media further supported the results of my genetic data. If BioI, a cytochrome P450, was important for biotin synthesis *B. subtilis* would depend on oxygen for growth in biotin-free media. I hypothesize that acquiring *bioW* and making *bioI* extraneous during the course of evolution might have enabled *B. subtilis* to survive in environments where oxygen is scarce and other electron acceptors are available. If so, *bioW* acquisition seems likely to have been a recent occurrence because the BioI sequence has not degenerated into an inactive protein. Consistent with this notion *bioW* is reported to have its own dedicated transcript in addition to the transcript that encodes bioW and other

genes of the operon (104) It should be noted that other bacteria that encode a BioW lack a BioI. Some of these are *Bacillus* (now *Lysinibacillus*) *sphaericus, S. aureus, Aquifex aeolicus* and *Chlamydia trachomatis.* The methanogenic archaeon, *Methanococcus jannaschii* also encodes a BioW but lacks BioI consistent with it strictly anaerobic lifestyle.

The mechanism of the *B. subtilis* pimelate synthesis remains to be determined. The most straightforward pathway would be if the fatty acid synthetic pathway could accept a malonyl-thioester as a primer in place of the thioesters of the usual acetyl or branched chain precursors. If so, when the pimelate chain was complete, it would be cleaved from ACP by a (presumably) specific pimeloyl-ACP thioesterase and BioW would convert the free pimelate to pimeloyl-CoA. Work to be published elsewhere argues that pimeloyl-CoA but not pimeloyl-ACP is a substrate for *B. subtilis* BioF. This pathway would be wasteful of ATP since an already activated pimeloyl carboxyl is cleaved and then reactivated, but since biotin synthesis is a very low demand pathway (36) there should be negligible effects on *B. subtilis* physiology.

| | | |
|---|---|---|
| ***B. subtilis*** | | |
| 168 | Wild type | (132) |
| MM43 | Δ*bioW* of 168, downstream reading promoter | This work |
| MM57 | Δ*bioI* of 168 | This work |
| MM59 | MM43 with ectopic *bioW* | This work |
| PY79 | Wild type | (132) |
| BI274 | PY79 *bio* operon driven by strong constitutive SPO1-26 promoter | (11, 104) |
| BI304 | PY79 *bio* operon driven by strong constitutive promoter SPO1-15 and containing SPO1-15 in place of the terminator upstream of *bioI* [$P_{15}$ *bioWAFDB (Δt)* $P_{15}$ *bioI]* | (104) |
| MM194 | Δ*lipM* of BI274 | This work |
| Cer-20 | Cerulenin-resistant derivative of 168 | (115) |
| NM57 | Δ*lipM*::Km | (84) |
| ***E. coli*** | | |
| NRD25 | Δ*(bioABFCD)::Cm* | (15) |
| ER90 | Δ*bioF bioC bioD*[a] | (14) |
| STL25 | Δ*bioC* Δ*bioH* | (75) |

**Table 3.1:** *E. coli* and *B. subtilis* strains.

a.The Δ*bioF* lesion is polar on *bioC* and *bioD*

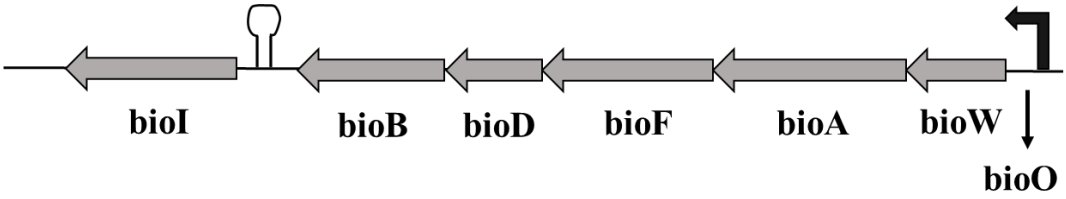



**Figure 3.1:** The *B. subtilis bio* operon consists of *bioW, bioA, bioF, bioD, bioB* and *bioI.* The arrow at the right indicates the promoter while the hairpin structure denotes the terminator upstream of *bioI* (104). The *bio* genes are transcribed from a promoter that is regulated in a biotin-dependent manner by the BirA, biotin protein ligase/repressor which binds to *bioO*, the operator sequence located upstream of *bioW.*

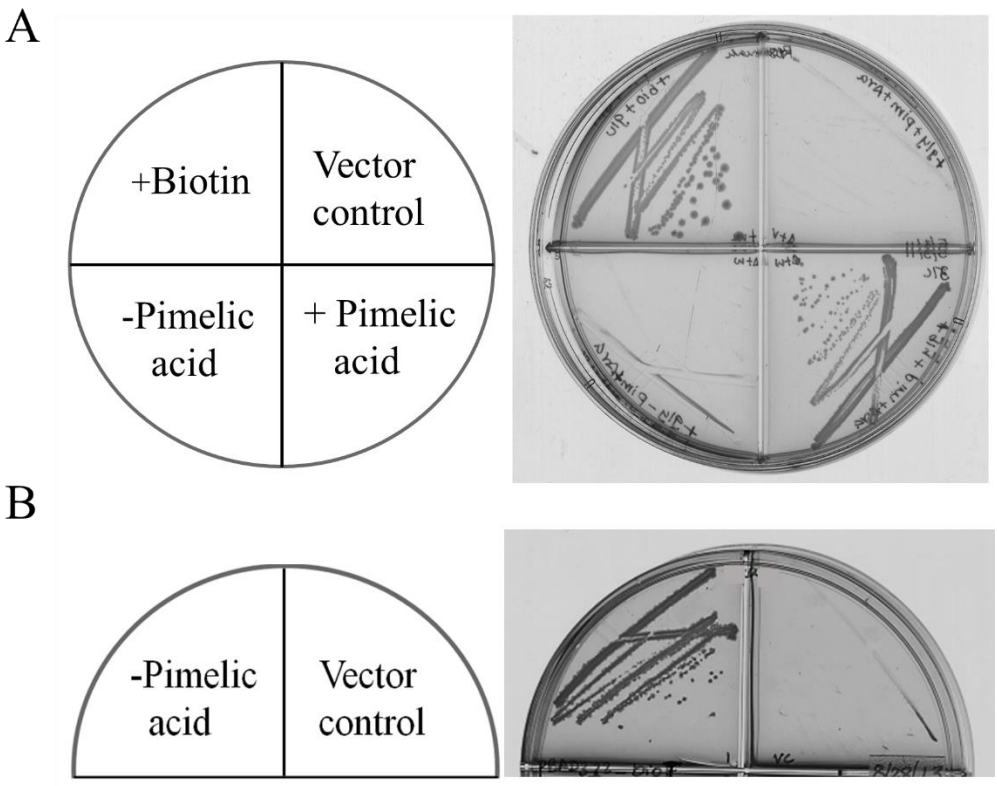



**Figure 3.2:** Complementation of the biotin requirement of the *E. coli* Δ*bioC* Δ*bioH* strain STL25 with *B. subtilis bioW* or *bioI.* A) *B. subtilis bioW* allowed growth of the *E. coli* Δ*bioC* Δ*bioH* strain in biotin-free minimal media only upon supplementation with pimelic acid. Expression of *bioW* was induced from the $P_{BAD}$ promoter of vector pBAD322 by arabinose. Vector control indicates the empty vector induced with arabinose. Pimelic acid (0.1 mM) was used for growth supplementation. B) *B. subtilis bioI* allowed growth of the *E. coli* Δ*bioC bioH* strain without supplementation with either pimelic acid or arabinose. *B. subtilis bioI* was inserted into vector pBAD322.

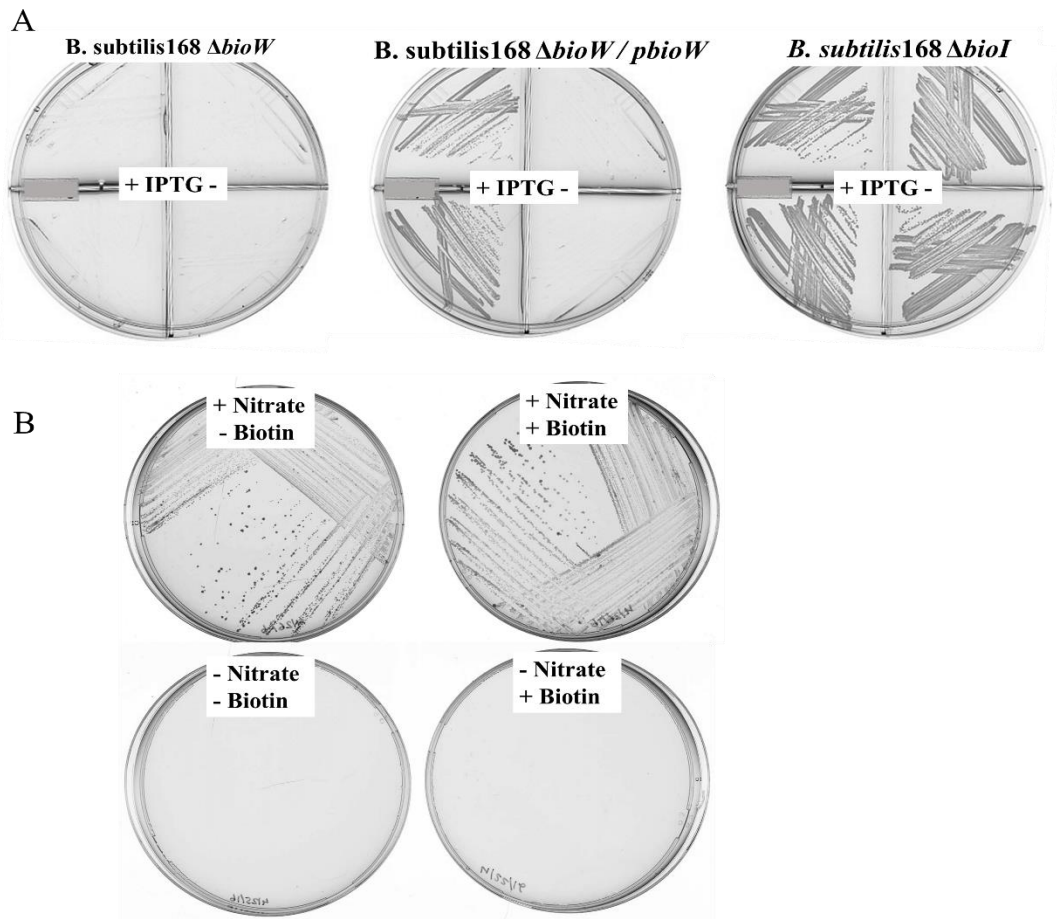



**Figure 3.3:** Growth phenotypes of *B. subtilis bioW* and *bioI* mutant strains. A) Phenotypes of *B. subtilis bioW* and *bioI* deletion mutants. Deletion of the chromosomal *bioW* through single crossover recombination by integration of recombinant vector pMUTIN4 blocked growth in biotin-free minimal media. Two separate colonies were streaked into upper and lower sectors of the plates sectioned by plastic walls. IPTG (1 mM) was used (left sectors, note the plus signs) to induce the expression of downstream genes whereas no IPTG was present in the right-hand sectors (minus signs). Expression of *bioW* from the Phyper-spank promoter of vector pDR111 inserted at an ectopic site (the *amyE* locus) restored growth only when promoter activity was induced with IPTG. No change in growth was observed when the chromosomal *bioI* gene was partially deleted and disrupted by insertion of the kanamycin cassette of plasmid pDG780. The *bioW* strain was MM43 whereas the Δ*bioI* strain was MM57. The derivative of MM43 with an ectopic *bioW* was strain MM59 (104). B) Anaerobic growth of *B. subtilis*. *B. subtilis* 168 was streaked and incubated on biotin-free minimal medium plates with nitrate as the terminal electron acceptor using the BD Gaspak EZ Anaerobe Pouch System. Growth was dependent on nitrate but not biotin indicating that *B. subtilis* synthesizes biotin anaerobically and therefore the oxygen-dependent cytochrome P450 enzyme, BioI is not required. The nitrate-dependent growth indicates that the conditions were anaerobic.

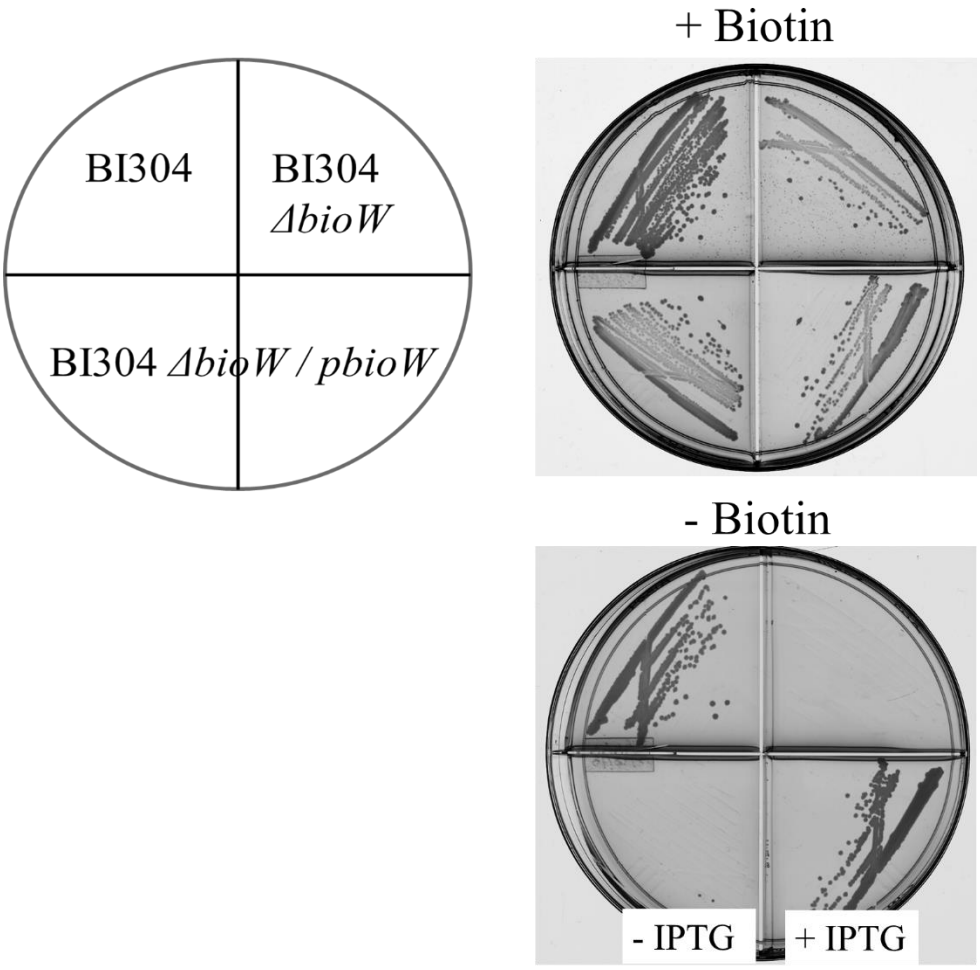



**Figure 3.4:** BI304 *ΔbioW* is unable to grow in biotin-free minimal media. BI304 is a strain in which the terminator preceding *bioI* is replaced with a constitutive promoter to allow overexpression of *bioI* (11). Top panel shows growth of all strains, indicated in the schematic, in biotin containing minimal media. Bottom panel shows growth of BI304 (wild-type) in biotin-free minimal media whereas BI304 *ΔbioW/ pbioW* (wild-type *bioW* integrated into an ectopic *amyE* site) grows only in presence of IPTG, which induces expression of ectopic *bioW*.

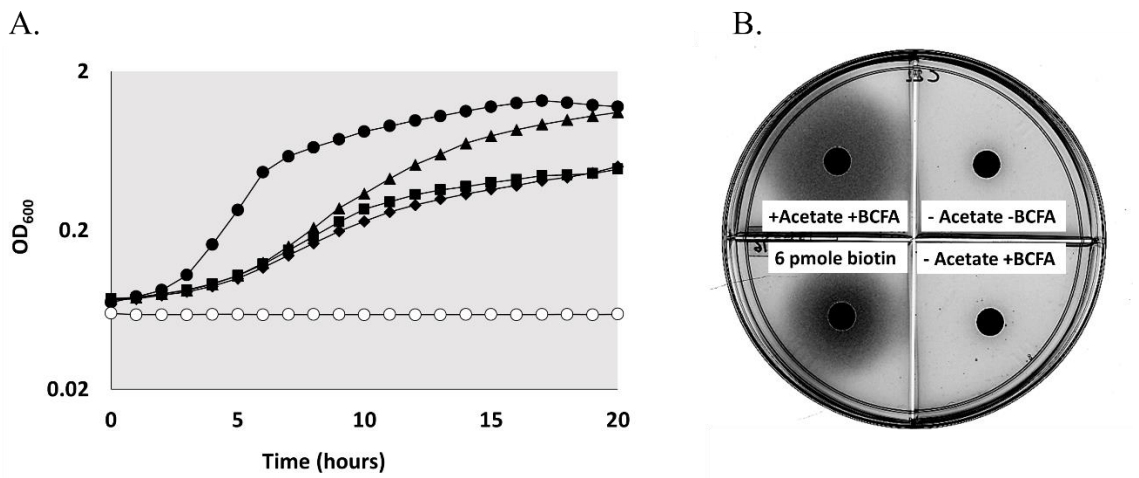



**Figure 3.5:** Acetate supplementation is required for growth and biotin production by strain MM194. A) Growth of strain MM194 with 1 mM sodium acetate and 0.1 mM each of the branched-chain fatty acid (BCFA) precursors isovaleric acid, isobutyric acid and 2-methylbutyric acid. Strains BI274 and MM194 were grown overnight in minimal medium supplemented with yeast extract. Cells were harvested, washed twice in Spizizen salts and subcultured in a 96 well plate to a final $OD_{600}$ of 0.012 in 300 µL of minimal media containing different growth supplements: (solid circles) BI274 without sodium acetate and BCFA precursors, (triangles) MM194 with sodium acetate and BCFA precursors, (squares) MM194 with BCFA precursors only, (diamonds) MM194 without acetate and BCFA precursors and (open circles) minimal media lacking cells. Strain BI274 grew without acetate and BCFA precursors whereas strain MM194 required both acetate and BCFA precursors for growth. B) Bioassay of biotin produced by strain MM194 plus acetate or BCFA precursors or both. Supernatants from the cultures were collected after growth on acetate and BCFA precursors and 10 µl of each was spotted into sterile paper disks on minimal agar plates containing the *E. coli* biotin auxotroph NRD25 (Δ*bioABFCD::Cm*). Production of biotin with acetate and BCFA precursors supplementation agrees with the growth dependence on these substrates (panel A). Known biotin amounts were used to quantitate biotin in each batch of the minimal plates used for bioassay.

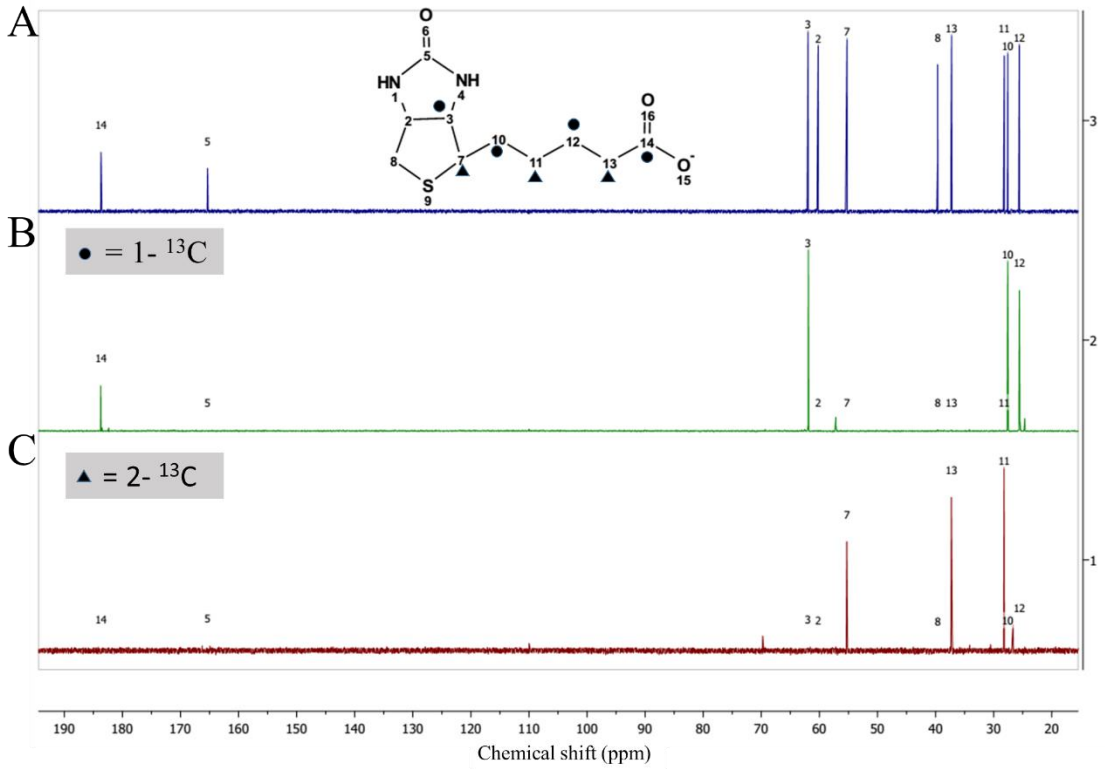



**Figure 3.6:** $^{13}$C-NMR spectrum of biotin from [1-$^{13}$C]acetate or [2-$^{13}$C]acetate. A) Natural abundance $^{13}$C spectrum of commercial biotin gives peaks with chemical shifts corresponding to each of the biotin carbon atoms(55, 113). Note that compared to my NMR program, Ifuku and Sanyal each use different biotin atom numbering systems. My C14 is C1 of Ifuku and C10 of Sanyal whereas my C3 is C7 of Ifuku and C3 of Sanyal. B) Spectrum of $^{13}$C-enriched carbons from [1-$^{13}$C]acetate. Carbons 3, 10, 12 and 14 of biotin (solid circles) are enriched with $^{13}$C. Carbon atoms 3 and 14 are derived from the carboxyl carbons of pimelate. C) Spectrum of $^{13}$C-enriched carbons from [2-$^{13}$C]acetate. Carbons 7, 11 and 13 of biotin (solid triangles) are enriched with $^{13}$C.

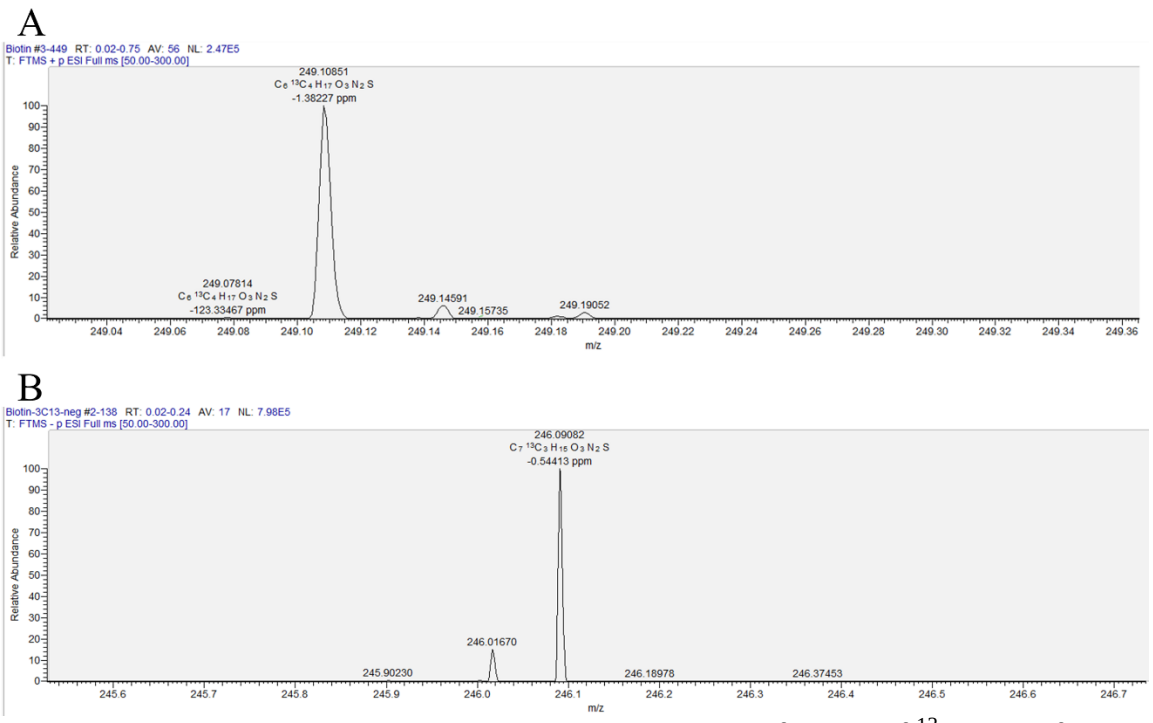



**Figure 3.7:** Electrospray ionization mass spectroscopy identification of $^{13}$C-biotin from A) [1-$^{13}$C]acetate (calculated mass of 249.109) in the positive ESI mode and B) [2-$^{13}$C]acetate (calculated mass 246.090) in the negative ESI mode. The monoisotopic mass of biotin is 244.088.

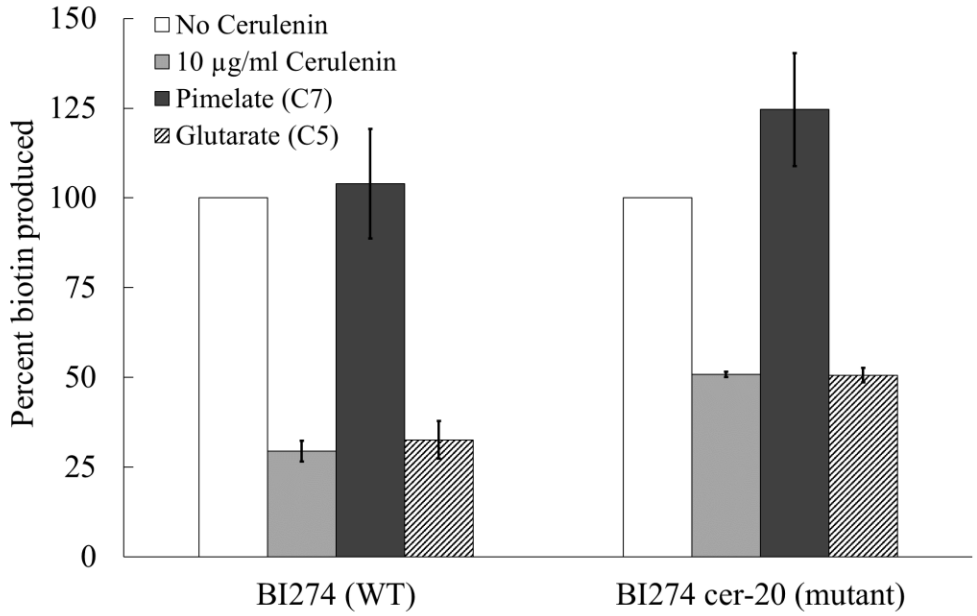



**Figure 3.8:** Biotin production by untreated (white bars) and cerulenin treated (light gray bars) resting cell cultures of BI274 (WT) and BI274 cer-20 (cerulenin-resistant mutant) strains. The cells were grown in 400 mL of minimal media overnight. The cells were harvested, washed and transferred to nitrogen-limited medium in 25 mL culture volumes and allowed to incubate for 30 min at 37°C before treatment with 10 µg/ml of cerulenin (light gray bars). Samples (2 mL) were taken at the end of every hour up to 6 h and either 0.1 mM pimelate (dark gray bars) or 0.1 mM glutarate (stippled bars) was added to the cerulenin-treated cultures and incubated for one additional hour before taking the final sample. Supernatants were harvested by centrifugation and filtered through a 0.22 µm membrane. Ten microliters of each sample was spotted onto paper disks placed on top of minimal medium plates containing a biotin auxotrophic strain of *E. coli* and a redox indicator (2,3,5-triphenyltetrazolium chloride) as in Figure 3-5. The percent of biotin production was calculated by comparing the amount of biotin produced with that of the untreated culture. The error bars represent standard errors. In the seven repetitions of this experiment, 100% biotin production ranged from 6 to 25 pmol in the bioassays. The final biotin measurements for the antibiotic-treated cultures were determined after 6 h in each experiment.

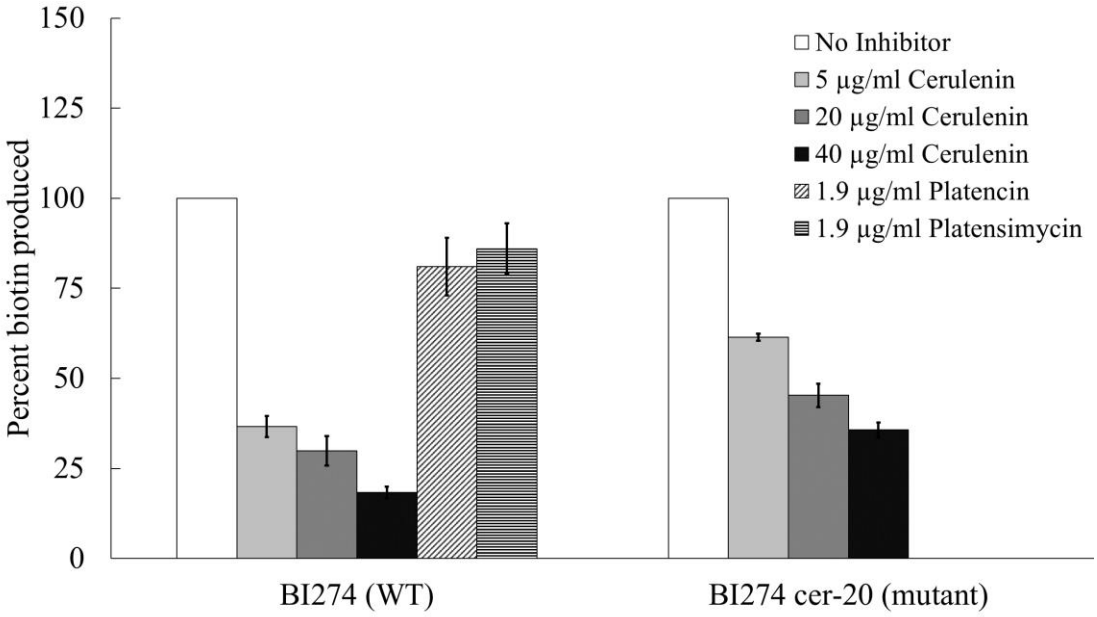



**Figure 3.9:** Biotin production by resting cells of the wild type strain BI274 and BI274 cer-20 its FabF (I108F) derivative. The methods were the same as in Figure 3-8 and Methods and Materials. Resting cells were treated with 5, 20 or 40 µg/ml of cerulenin in both panels whereas (left panel only) either 1.9 µg/ml of platencin or 1.9 µg/ml of platensimycin was added. The biotin contents of the supernatants were determined by bioassay. The error bars represent standard errors.

# CHAPTER 4
# CONCLUSIONS

## SUMMARY OF FINDINGS

The specific aim of this Thesis is to characterize the biotin biosynthesis pathway of *Bacillus subtilis.* Numerous studies in the past have successfully illustrated the mechanism of second half of the pathway that forms the two rings of the biotin molecule. However, our knowledge about the initial stage of biotin synthesis remained incomplete. Several conjectures existed to predict the synthesis of a pimelate (a seven-carbon α, ω-dicarboxylic acid) moiety intermediate that separated the pathway into two distinct halves. The pimelate moiety was believed to be synthesized from acetate with CoA as the carrier molecule. In 2010, Lin et. al. demonstrated a mechanism used by *Escherichia coli* to synthesize its pimeloyl-ACP (thioester linkage with acyl carrier protein) intermediate by hijacking the fatty acid synthesis pathway. A SAM-dependent methyl transferase, BioC methylates malonyl-ACP to allow its entry into the fatty acid synthesis pathway by shielding the carboxyl group of the malonyl moiety. The malonyl moiety converts into pimeloyl moiety in two rounds of reduction, dehydration and reduction. BioH, a carboxylesterase cleaves the methyl ester to generate pimeloyl-ACP for the second stage of biotin synthesis. This mechanism of hijacking the fatty acid synthesis pathway for biotin production is present in organisms that contain *bioC* and *bioH* or their homologs. However, we know that there are several other organisms that do not have genes homologous to *bioC* and *bioH.* I chose *B. subtilis* has a model organism not only to study the other possible mechanisms of pimeloyl moiety generation but also to investigate the curious redundancy of the genes involved in pimelate moiety generation: *bioI* and *bioW.*

In Chapter 2, I describe the unique characteristics of BioW as a pimeloyl-CoA synthetase. BioW is an atypical acyl-CoA ligase/synthetase capable in the ATP-dependent conversion of pimelic acid to pimeloyl-CoA in a two-step reaction involving the formation of a pimeloyl-AMP intermediate (Figure 2-1A). I began characterizing BioW in the beginning of my Thesis work because the protein sequence had no

similarities with known acyl-CoA ligases and it was easily purified in large amounts. I discovered a unique function of BioW during enzymatic analysis. Apart from having the expected function of generating pimeloyl-CoA, BioW can also proofread noncognate substrates to ensure proper chain length incorporation into the biotin molecule. Noncognate substrates like glutaric acid (five-carbon α, ω-dicarboxylic acid, C5) and suberic acid (an eight-carbon α, ω-dicarboxylic acid, C8) can be eliminated by hydrolysis of the acyl-adenylate intermediate. The production of AMP (end product) in the absence of CoA (acceptor of acyl intermediate) provided a strong evidence for this conclusion. This is similar to pretransfer editing of aminoacyl-tRNA synthetase to choose the cognate amino acid from another similar in structure. These noncognate intermediates either hydrolyze spontaneously in solution after being removed from the enzyme or are hydrolyzed in the enzyme active site. I performed pH stability experiments to show that the hydrolysis takes place in the enzyme active site. Acyl-adenylates are prone to hydrolysis at higher pH. Increasing the solvent pH did not have an effect in the hydrolysis of noncognate acyl-adenylate intermediates. However, increasing the solvent pH increased the rate of hydrolysis of the cognate pimeloyl-AMP intermediate indicating the accessibility of solvent into the active site upon binding of cognate substrate. In absence of CoA, the rate of AMP production was linear with time in case of glutaric and suberic acid whereas pimeloyl-AMP saturated the enzyme active site. In conclusion, I showcased a unique proofreading ability of an uncommon acyl-CoA ligase, BioW.

Despite the lack of evidence to show a link between BioW and aminoacyl-tRNA synthetases, this showed how *B. subtilis* avoids the expensive route to making an unusable derivative of biotin. Collaboration with Dr. Satish Nair and his group led to the crystallization of *Aquifex aeolicus* BioW that revealed a unique crystal structure of the protein (33). The structure contains a smaller N-terminal domain and a larger C-terminal domain with a modified Rossman fold common to nucleotide binding proteins (112). AaBioW also revealed proofreading activity of a noncognate substrate (C5) verifying the unique secondary function of an acyl-CoA ligase. This work made it possible to look at active site residues that interact with the substrates as well as test their role in the proofreading activity. Analysis on one of the AaBioW mutants provided evidence for the

role of single residue (Arg159) in the recognition of the cognate substrate. Wang and workers also recently illustrated the crystal structure of *B. subtilis* BioW that resembles the structure of *A. aeolicus* BioW (127).

All prior studies characterized BioW as a pimeloyl-CoA synthetase. However, its role in the pathway was undefined given the presence of BioI, speculated as a source of pimeloyl-ACP. I started with a hypothesis that BioI is the major enzyme that supplies pimeloyl-ACP for biotin synthesis. This is because BioI is a well characterized cytochrome P450 with a crystal structure that clearly outlines the binding pocket where oxidation of long chain fatty acyl chains takes place to generate pimeloyl-ACP. However, previous studies showed genetic knockout of *bioW* caused biotin auxotrophy while removal of *bioI* merely caused a bradytrophic phenotype in *B. subtilis* PY79 prototrophic strain. In Chapter 3, I carried out *bioW* gene disruption by inserting plasmid pMUTIN4 with an inducible promoter for the expression of downstream genes, to avoid polarity effects. This clearly caused biotin auxotrophy as previously described by Bower and coworkers (10). Ectopic expression of the wild-type *bioW* allowed restoration of growth in the *bioW* mutant. However, *bioI* disruption by insertion of kanamycin cassette did not have an effect on the growth of *B. subtilis* 168 in biotin-free minimal media. Additionally, I observed growth of *B. subtilis* 168 in anaerobic minimal media containing nitrate as the terminal electron acceptor, regardless of the presence of biotin. Since oxygen is required for a cytochrome P450 function, inactivation of BioI in anaerobic environment should have prevented *B. subtilis* growth. The insignificance of BioI in *B. subtilis* biotin synthesis is curious because *bioI* complements an *E. coli ΔbioC ΔbioH* mutant strain deficient in generating pimeloyl-ACP very well. The most straightforward explanation is that BioI might be a remnant of an oxygen-dependent strain of *B. subtilis*. Evolution may have favored acquisition of BioW to allow the organism to continue surviving in oxygen scarce environments.

The importance of pimeloyl-CoA as the intermediate in biotin synthesis is established because of the significance of BioW. The question was where does BioW acquire free pimelate? Pimelate is not a common molecule and is not known to exist in

central metabolism. I attempted to decipher the source of pimelate by looking into genomic libraries of *B. subtilis* that could complement *E. coli* mutants for the lack of pimelate. I randomly mutagenized *B. subtilis* strain to uncover genetic elements responsible for pimelate synthesis. After years of failed attempts to discover the source of pimelate through genetic approaches, I concentrated my efforts to a different methodology- $^{13}$C-NMR.

This approach is based on a previous study by Sanyal et al. (113) and Ifuku et al. (55) where they determined the origin of all carbons of the biotin molecule in *E. coli* from [1-$^{13}$C]- or [2-$^{13}$C]-labeled acetate. Since the labeling pattern in biotin was different from when *E. coli* cells were grown in [1-$^{13}$C]acetate compared to [2-$^{13}$C]acetate, they independently concluded that the pimelate moiety originating from acetate units has to be attached to a fixed end via a thioester linkage; thus, *E. coli* cannot use free pimelate for biotin synthesis. Pimelate is a symmetrical molecule which should incorporate into biotin resulting in a non-specific labeling pattern. *B. subtilis* on the other hand can use free pimelate for biotin synthesis so, this approach could be used not only to provide evidence for the presence of free pimelate in the cells but also as a guide to a pathway for pimelate synthesis.

Maximum incorporation of $^{13}$C-acetate is essential to harvest pure highly enriched $^{13}$C-biotin for NMR analysis. *B. subtilis*, unlike *E. coli,* cannot use acetate as sole carbon source due to lack of the glyoxylate cycle. I acquired a biotin overproducing strain (BI274) from a previous work that is patented by Bower and coworkers (11). I introduced a *ΔlipM* mutation into the strain to inactivate the lipoic acid requiring pyruvate dehydrogenase to block the supply of acetate from metabolized glucose used in the minimal media. The LipM mutation blocks lipoic acid synthesis required for pyruvate dehydrogenase and branched-chain dehydrogenase activity. Hence, the newly engineered strain could now use $^{13}$C-acetate and branched chain fatty acid precursors for growth in minimal media. Strain BI274 *ΔlipM* was grown in [1-$^{13}$C]acetate or [2-$^{13}$C]acetate in glucose containing minimal media. Twenty liters of *B. subtilis* culture was used to extract biotin in several steps of purification. The extract always consisted of a thick brown

coloring which was difficult to remove. Finally, affinity purification with a mutant derivative of shwanavidin, an avidin from *Shewanella denitrificans,* removed most of the colored contaminant. Biotin eluted with heated water was concentrated and dissolved in deuterium for NMR analysis. This protocol required several optimizations due to the necessity to keep biotin from contamination from organic solvents, which are detectable in $^{13}$C-NMR if present in abundant amounts.

As expected, the label from [1-$^{13}$C]acetate was incorporated into both terminal carbons arising from the pimelate molecule confirming the presence of the pimelate in *B. subtilis.* Additionally, the alternate labeling pattern of the valeric acid carbons from [1-$^{13}$C]acetate and [2-$^{13}$C]acetate fed cultures indicated the involvement of a pathway similar to fatty acid synthesis for the synthesis of pimelate. Following up on this result, I tested the effect of several fatty acid enzyme inhibitors on biotin synthesis via a resting cell bioassay. If fatty acid synthesis is involved in supplying pimelate for biotin synthesis, blocking it with its inhibitors should have a negative effect on the production of biotin. I used resting cells of *B. subtilis* for this experiment to normalize the effect of cell growth in both the wildtype and mutant cells. Administration of cerulenin, a FabF inhibitor, had significant negative effect on biotin synthesis as seen in the bioassay media containing *E. coli* biotin auxotrophic strain. However, biotin production in a cerulenin resistant derivative was not as severely affected as in the wild-type strain. Additionally, supplementation of pimelate, but not glutarate, into the cerulenin treated cells restored the production of biotin to normal levels. Other inhibitors, platencin and platensimycin were not as effective in causing inhibition of biotin synthesis as was cerulenin.

In summary, I have characterized one of the main enzymes that drives the initial step of biotin synthesis in *B. subtilis* and answered a long-standing question about the origin of pimelic acid. BioW is a pimeloyl-CoA synthetase which converts free pimelic acid to pimeloyl-CoA and proofreads noncognate dicarboxylic acid by hydrolyzing the reaction intermediate. This is the first demonstration of an acyl-CoA ligase having a function similar to an aminoacyl-tRNA synthetase. BioI, on the other hand, is not essential for biotin synthesis and appears to be an evolutionary remnant. The substrate of

BioW, pimelic acid has its origins linked to fatty acid synthesis. Although this link is analogous to the genesis of pimeloyl-ACP from fatty acid synthesis in *E. coli,* further details need to be uncovered to determine the true pathway for pimelate synthesis in *B. subtilis.*

## FUTURE DIRECTIONS

### Identification of preferred substrate of *B. subtilis* BioF: pimeloyl-CoA or pimeloyl-ACP

I briefly introduced BioF in Chapter 1 as a 8-amino-7-oxononanoate synthase that carry out decarboxylative condensations between amino acids and carboxylic acid thioesters to form 8-amino-7-oxononanoate. In general, BioF is the enzyme that initiates the second stage of the pathway by utilizing the pimelate thioester intermediate (Figure 1-2). Although I have shown BioI as a non-essential enzyme that produces pimeloyl-ACP, it would be interesting to compare the activity of *B. subtilis* BioF with pimeloyl-CoA and pimeloyl-ACP.

The crystal structure of *E. coli* BioF is shaped like a left-hand palm with a deep cleft for substrate binding between the "thumb" and the "fingers" (2). Alexeev and coworkers speculated the cluster of positively charged arginine residues could be favorable binding site for phosphate of pimeloyl-CoA (Figure 4-1A, 4-2). However, they were unable to soak pimeloyl-CoA into the crystals to obtain a substrate-bound crystal structure. Obviously, this speculation was based on studies prior to the work of Lin et al., when pimeloyl-CoA, not pimeloyl-ACP, was thought to be the pimelate thioester intermediate. In hindsight, I can now predict that these positively charged residues might play a role in binding the acidic ACP protein. Although the activity of *E. coli* BioF has not been tested with pimeloyl-ACP to confirm this hypothesis, it is clear that *E. coli* BioF should prefer pimeloyl-ACP. I have shown the activity of *E. coli* BioF with pimeloyl-ACP (Figure 4-3C) in a gel-shift experiment to show the conversion of pimeloyl-ACP to holo-ACP. Webster et al. (128) compared enzyme kinetics of *E.coli* BioF and *B. sphaericus* BioF with pimeloyl-CoA. They showed that the $K_m$ of *E. coli* BioF for

pimeloyl-CoA is 16-fold higher than that of the *B. sphaericus* BioF indicating that the *E. coli* enzyme has a much lower specificity for pimeloyl-CoA.

The modeled structure of *B. subtilis* BioF is comparable to that of *E. coli* BioF. However, *Bs*BioF model structure reveals interesting differences in key residues between the two proteins. The positively charged patch populated by arginine residues in *E. coli* BioF is replaced by polar uncharged glutamine or non-polar valine residues (Figure 4-2). This might be because *B. subtilis* BioF prefers a different substrate, pimeloyl-CoA, which may not require multiple charged residues for proper binding. *E. coli* BioF active site suited for ACP-bound substrate might be large enough to admit a CoA derivative since both ACP and CoA share the phosphopantetheine group. However, in case of *B. subtilis* BioF, the substrate binding site suited for a CoA-bound substrate may be too small to accept an ACP-bound substrate. It would be interesting to test this hypothesis by conducting site-directed mutagenesis replacing the native residues with either arginine or alanine to test for activity with pimeloyl-ACP (arginine substitutions) or loss of activity with pimeloyl-CoA (alanine substitutions).

I have conducted some preliminary experiments to characterize *Bs*BioF's activity and its preference for pimeloyl-CoA. The protein was purified under denaturing conditions due to insolubility upon overexpression in *E. coli* (Figure 4-3A). Refolded *Bs*BioF showed activity comparable to native *B. sphaericus* BioF, with pimeloyl-CoA (Figure 4-3B) but not with pimeloyl-ACP (Figure 4-3C). This clearly indicates that *B. subtilis* BioF prefers pimeloyl-CoA in *in vitro* reactions. I carried out an *in vivo* test to validate this result. *E. coli* BioF can complement *B. subtilis* Δ*bioW* mutant but not Δ*bioW* Δ*bioI* mutant. The result clearly indicates that the product of BioI, pimeloyl-ACP, is used by *E. coli* BioF to restore the growth by bypassing the BioW reaction. Note that *B. subtilis* BioF cannot restore the growth in a Δ*bioW* mutant strain despite having a functional BioI that can produce pimeloyl-ACP. Further experiments can be done to analyze KAPA formation from reactions with either pimeloyl-CoA or pimeloyl-ACP to compare the activity of BioF with the two substrates. One could do that by conducting a bioassay on an *E. coli* BioF test strain. Another method would be to derivatize KAPA

with chemicals such as fluorescent dansyl-chloride to quantitate the amount of KAPA formed through LC/GC-MS.

FIGURES

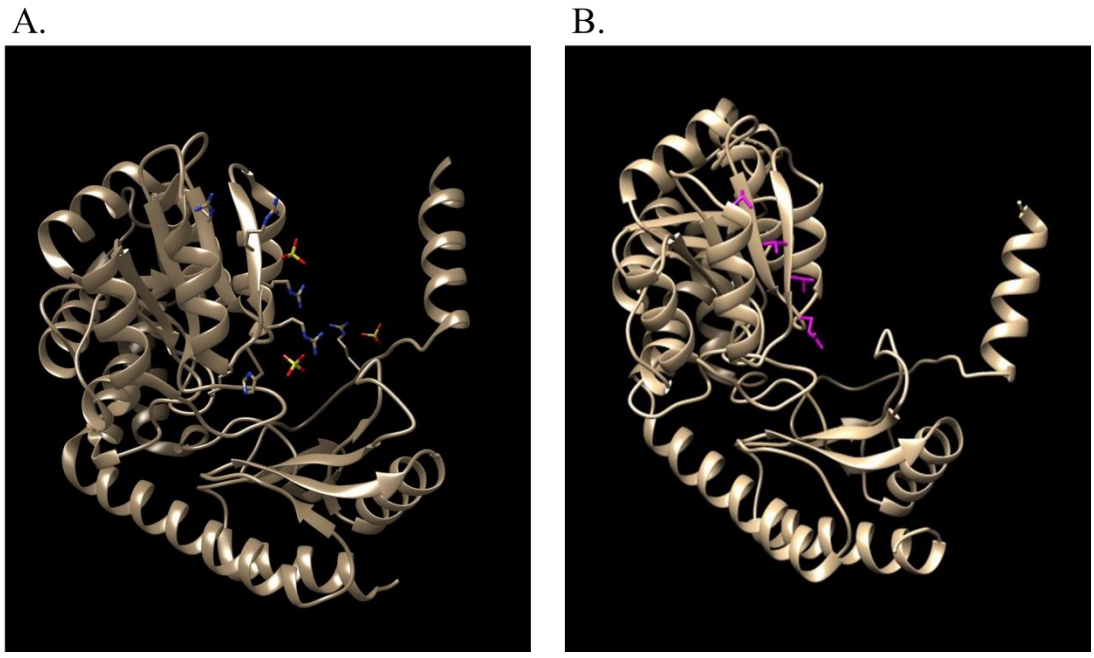



**Figure 4.1:** Crystal structures of *E. coli* BioF (left panel) and *B. subtilis* BioF (right panel). A) *E. coli* BioF is shaped like a left-hand palm with a deep cleft. The "thumb" is the N-terminal, the "fingers" make up the central domain, and the base of the palm is the C-terminal domain. A positive patch made up of arginine residues, represented in blue, is predicted to be the binding site of CoA. B) The model of *B. subtilis* BioF resembles the *E. coli* BioF structure. However, the arginine residues are replaced by uncharged glutamine or non-polar valine residues, represented in magenta. The image of *E. coli* BioF (1DJE) was downloaded from Protein Data Bank. *B. subtilis* BioF was modeled using Chimera.

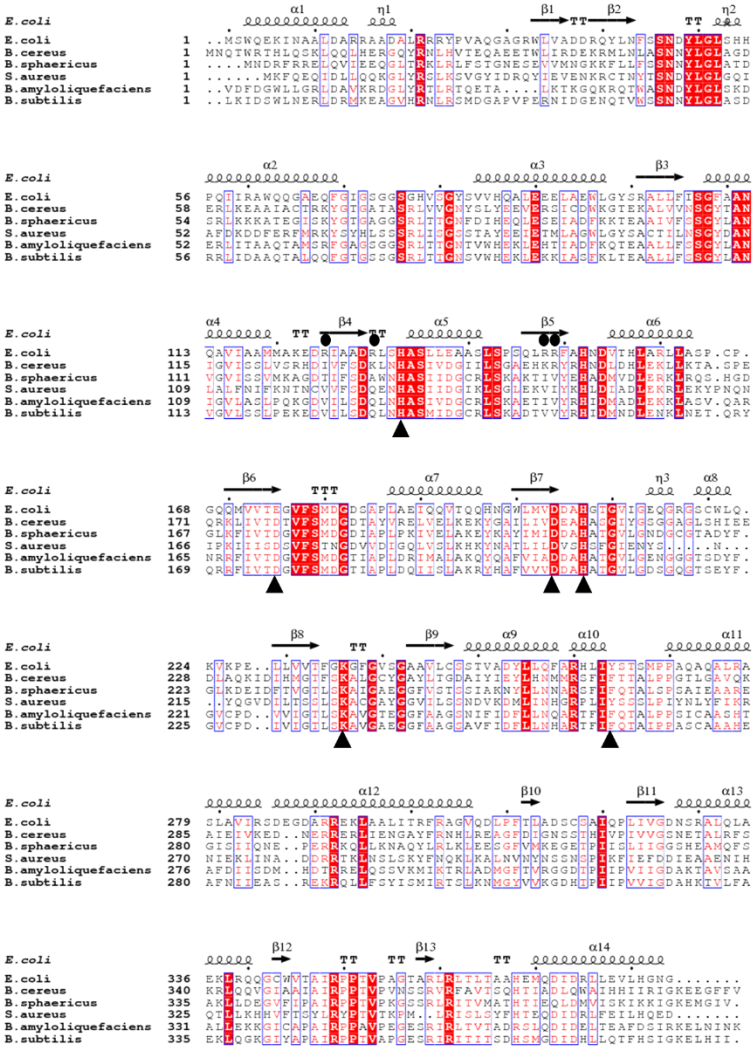



**Figure 4.2:** Amino acid sequence alignments of BioF from *E. coli*, *Bacillus cereus*, *Bacillus* (now *Lysinibacillus) sphaericus*, *Staphylococcus aureus*, *Bacillus amyloliquefaciens* and *B. subtilis* using the amino acid sequence of *E. coli* BioF as template. The black triangles represent the conserved active sites whereas the black circles represent the arginine residues that contribute to a positive patch for substrate binding.

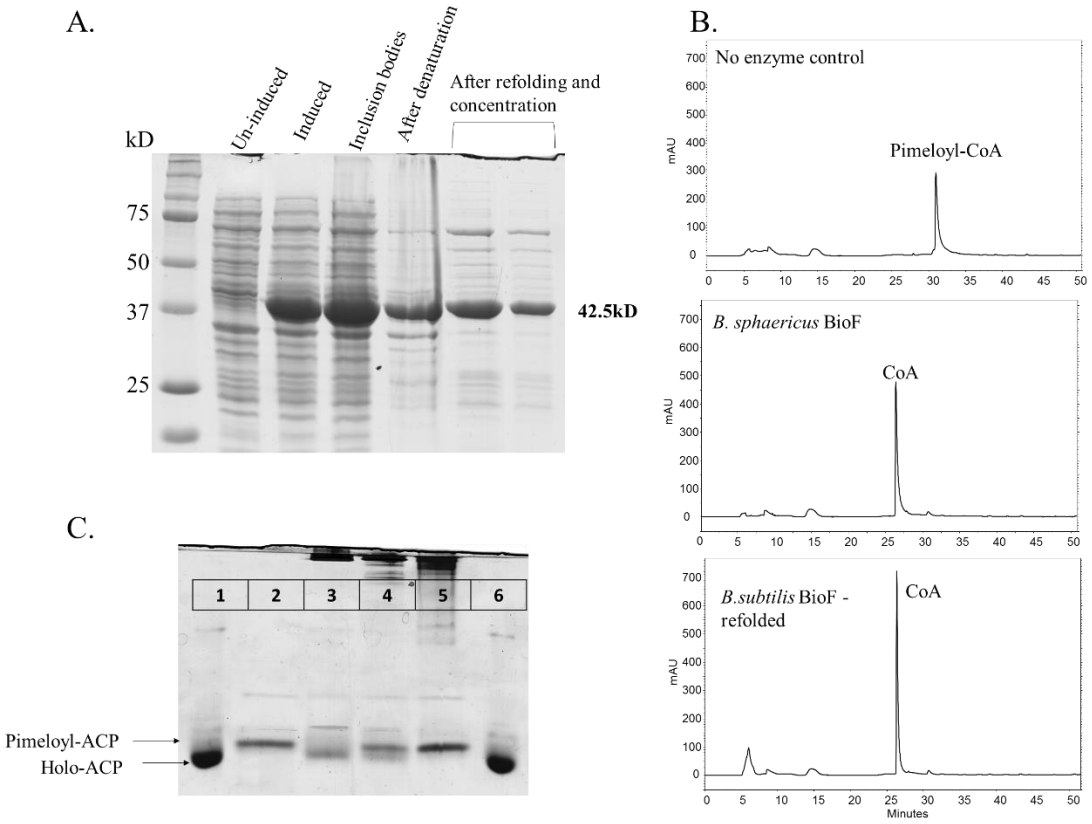



**Figure 4.3**: *B. subtilis* BioF enzymatic activity. A) SDS-PAGE gel showing the purification of *Bs*BioF under denaturing conditions. The molecular weight is 42.5kD. B) HPLC analyses of CoA formation from pimeloyl-CoA after reaction with *B. sphaericus* BioF (middle panel) or *B. subtilis* BioF (bottom panel) at 37°C for 1 hr. The top panel is a no enzyme control reaction showing the substrate, pimeloyl-CoA. C) Electrophoretic mobility shift assay showing reaction of *E. coli* BioF (lane 3), *B. sphaericus* BioF (lane 4) and *B. subtilis* BioF (lane 5) with pimeloyl-ACP as substrate to form holo-ACP. Lanes 1 and 6 holo-ACP controls and lane 2 is a no enzyme control showing pimeloyl-ACP.

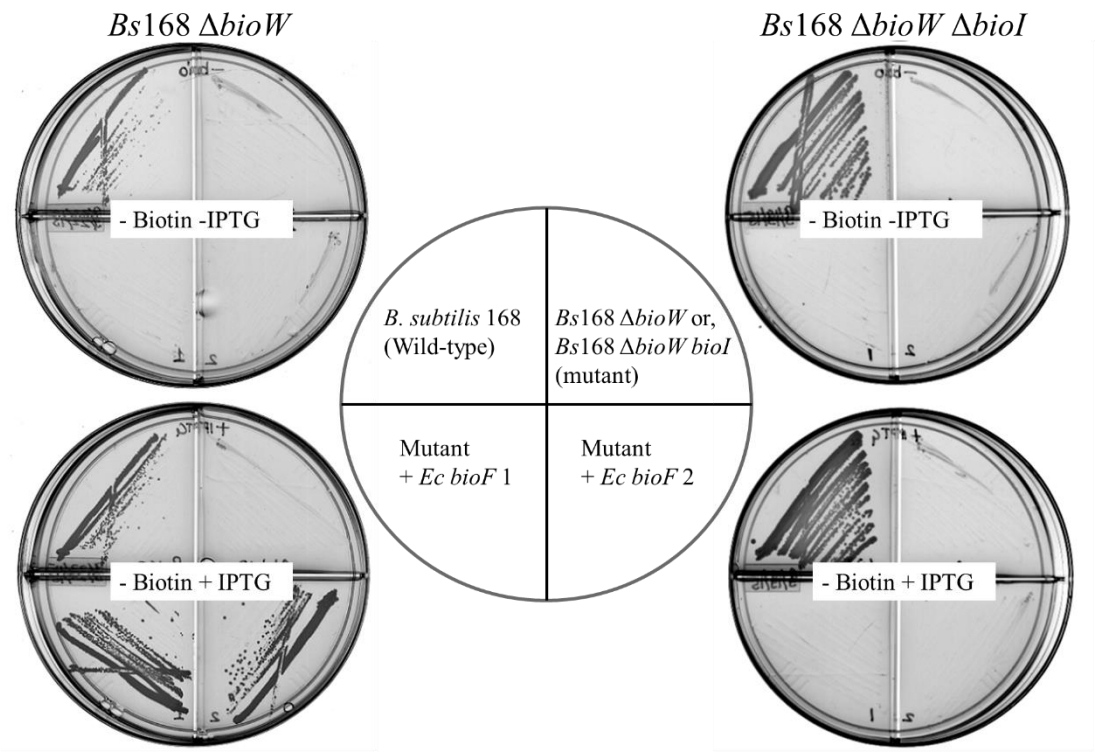



**Figure 4.4**: *In vivo* complementation assay in *B. subtilis* Δ*bioW* and Δ*bioW* Δ*bioI* mutants. The plates are sectioned into quarters by plastic walls. The schematic of the four quarter sections represent the strain streaked in each quarter. The plates contain biotin-free minimal media either with (bottom plates) or without (top plates) IPTG. Bottom sections of all the plates contain two separate colonies of the mutant strain with *E. coli bioF* integrated in an ectopic site (*amyE* locus) in the chromosome.

# CHAPTER 5

# REFERENCES


1. Agarwal V, Lin S, Lukk T, Nair SK, Cronan JE. 2012. Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. Proceedings of the National Academy of Sciences 109:17406-17411.

2. Alexeev D, Alexeeva M, Baxter RL, Campopiano DJ, Webster SP, Sawyer L. 1998. The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. Journal of Molecular Biology 284:401-419.

3. Allison FE, Hoover SR, Burk D. 1933. A respiratory coenzyme. Science 78:217-218.

4. Baldwin AN, Berg P. 1966. Transfer ribonucleic acid-induced hydrolysis of valyladenylate bound to isoleucyl ribonucleic acid synthetase. J Biol Chem 241:839-45.

5. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R. 2011. The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry 50:4402-4410.

6. Barker DF, Campbell AM. 1980. Use of bio-lac Fusion Strains to Study Regulation of Biotin Biosynthesis in Escherichia coli. Journal of Bacteriology 143:789-800.

7. Berkovitch F, Nicolet Y, Wan JT, Jarrett JT, Drennan CL. 2004. Crystal Structure of Biotin Synthase, an S-Adenosylmethionine-Dependent Radical Enzyme. Science 303:76-79.

8. Bi H, Zhu L, Jia J, Cronan JE. 2016. A Biotin Biosynthesis Gene Restricted to Helicobacter. Scientific Reports 6:21162-21174.

9. Birch TW, György P. 1939. Physiochemical properties of the factor (vitamin H) curative of egg white injury. Journal of Biological Chemistry 131:761-766.

10. Bower S, Perkins JB, Yocum RR, Howitt CL, Rahaim P, Pero J. 1996. Cloning, sequencing, and characterization of the Bacillus subtilis biotin biosynthetic operon. Journal of Bacteriology 178:4122-4130.

11. Bower SG, Perkins JB, Yocum RR, Pero JG. 2000. Biotin biosynthesis in Bacillus subtilis patent 6057136.

12. Brown SD, Ronson CW, Sullivan JT, Yocum RR. 2001. The bio operon on the acquired symbiosis island of Mesorhizobium sp. strain R7A includes a novel gene involved in pimeloyl-CoA synthesis. Microbiology 147:1315-1322.

13. Butterworth PH, Bloch K. 1970. Comparative aspects of fatty acid synthesis in *Bacillus subtilis* and *Escherichia coli*. Eur J Biochem 12:496-501.

14. Choi-Rhee E, Cronan JE. 2005. Biotin Synthase Is Catalytic In Vivo, but Catalysis Engenders Destruction of the Protein. Chemistry & Biology 12:461-468.

15. Choi-Rhee E, Cronan JE. 2005. A Nucleosidase Required for In Vivo Function of the S-Adenosyl-L-Methionine Radical Enzyme, Biotin Synthase. Chemistry & Biology 12:589-593.

16. Christensen QH, Cronan JE. 2009. The *Thermoplasma acidophilum* LplA-LplB complex defines a new class of bipartite lipoate-protein ligases. J Biol Chem 284:21317-26.

17. Cleary PP, Campbell A. 1972. Deletion and complementation analysis of biotin gene cluster of Escherichia coli. Journal of Bacteriology 112:830-839.

18. Cronan JE. 2006. A family of arabinose-inducible Escherichia coli expression vectors having pBR322 copy control. Plasmid 55:152-157.

19. Cryle MJ, Schlichting I. 2008. Structural insights from a P450 Carrier Protein complex reveal how specificity is achieved in the P450BioI ACP complex. Proceedings of the National Academy of Sciences 105:15696-15701.

20. Del Campillo-Campbell a, Kayajanian G, Campbell a, Adhya S. 1967. Biotin-requiring mutants of Escherichia coli K-12. Journal of bacteriology 94:2065-6.

21. Demoss JA, Genuth SM, Novelli GD. 1956. The Enzymatic Activation of Amino Acids Via Their Acyl-Adenylate Derivatives. Proc Natl Acad Sci U S A 42:325-32.

22. DeTitta GT, Edmonds JW, Stallings W, Donohue J. 1976. Molecular structure of biotin. Results of two independent crystal structure investigations. Journal of the American Chemical Society 98:1920-1926.

23. Dey S, Lane JM, Lee RE, Rubin EJ, Sacchettini JC. 2010. Structural Characterization of the Mycobacterium tuberculosis Biotin Biosynthesis Enzymes 7,8-Diaminopelargonic Acid Synthase and Dethiobiotin Synthetase. Biochemistry 49:6746-6760.

24. Du Vigneaud V, Dittmer K, Hague E, Long B. 1942. The growth-stimulating effect of biotin for the Diphtheria bacillus in the absence of pimelic acid. Science 96:186-187.

25. Du Vigneaud V. 1942. The structure of biotin. Science 96:455-461.

26. Du Vigneaud V, Melville DB, György P, Rose CS. 1940. On the identity of vitamin H with biotin. Science 92:62-63.

27. Dubnau D, Davidoff-Abelson R. 1971. Fate of transforming DNA following uptake by competent Bacillus subtilis: I. Formation and properties of the donor-recipient complex. Journal of Molecular Biology 56:209-221.

28. Eakin RE, McKinley WA, Williams RJ. 1940. Egg-white injury in chicks and its relationship to a deficiency of vitamin H (biotin). Science 92:224-225.

29. Eisenberg MA, Star C. 1968. Synthesis of 7-oxo-8-aminopelargonic acid, a biotin vitamer, in cell-free extracts of Escherichia coli biotin auxotrophs. Journal of bacteriology 96:1291-7.

30. Eisenberg MA. 1973. Biotin: biogenesis, transport, and their regulation. Adv Enzymol Relat Areas Mol Biol 38:317-72.

31. Eisenberg MA. 1962. The incorporation of 1,7 C14 pimelic acid into biotin vitamers. Biochemical and Biophysical Research Communications 8:437-441.

32. Elford HL, Wright LD. 1963. The incorporation of pimelic acid as a unit in the biosynthesis of biotin. Biochemical and Biophysical Research Communications 10:373-378.

33. Estrada P, Manandhar M, Dong S-H, Deveryshetty J, Agarwal V, Cronan JE, Nair SK. 2017. The pimeloyl-CoA synthetase BioW defines a new fold for adenylate-forming enzymes. Nature Chemical Biology In press.

34. Fan S, Li D-F, Wang D-C, Fleming J, Zhang H, Zhou Y, Zhou L, Zhou J, Chen T, Chen G, Zhang X-E, Bi L. 2015. Structure and function of Mycobacterium smegmatis 7-keto-8-aminopelargonic acid (KAPA) synthase. International Journal of Biochemistry and Cell Biology 58:71-80.

35. Farrar CE, Siu KKW, Howell PL, Jarrett JT. 2010. Biotin Synthase Exhibits Burst Kinetics and Multiple Turnovers in the Absence of Inhibition by Products and Product-Related Biomolecules. Biochemistry 49:9985-9996.

36. Feng Y, Zhang H, Cronan JE. 2013. Profligate biotin synthesis in alpha-proteobacteria - a developing or degenerating regulatory system? Mol Microbiol 88:77-92.

37. Feng Y, Napier BA, Manandhar M, Henke SK, Weiss DS, Cronan JE. 2014. A Francisella virulence factor catalyses an essential reaction of biotin synthesis. Molecular Microbiology 91:300-314.

38. Fersht A. 1999. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding W. H. Freeman and Company, New York, NY.

39. Garfinkel D. 1958. Studies on pig liver microsomes I. Enzymic and pigment composition of different microsomal fractions. Archives of Biochemistry and Biophysics 77:493-509.

40. Gloeckler R, Ohsawa I, Speck D, Ledoux C, Bernard S, Zinsius M, Villeval D, Kisou T, Kamogawa K, Lemoine Y. 1990. Cloning and characterization of the Bacillus sphaericus genes controlling the bioconversion of pimelate into dethiobiotin. Gene 87:63-70.

41. Goldberg MW, Sternbach L. 1949. United States Patent 2, 489, 237.

42. Green AJ, Munro AW, Cheesman MR, Reid Ga, von Wachenfeldt C, Chapman SK. 2003. Expression, purification and characterisation of a Bacillus subtilis ferredoxin: a potential electron transfer donor to cytochrome P450 BioI. Journal of inorganic biochemistry 93:92-9.

43. Green AJ, Rivers SL, Cheesman M, Reid GA, Quaroni LG, Macdonald IDG, Chapman SK, Munro AW. 2001. Expression, purification and characterization of cytochrome P450 Biol: a novel P450 involved in biotin synthesis in Bacillus subtilis. JBIC Journal of Biological Inorganic Chemistry 6:523-533.

44. Guérout-Fleury AM, Shazand K, Frandsen N, Stragier P. 1995. Antibiotic-resistance cassettes for Bacillus subtilis. Gene 167:335-6.

45. Gulick AM. 2009. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. ACS Chem Biol 4:811-27.

46. György P, Melville DB, Burk D, Du Vigneaud V. 1940. The possible identity of vitamin H with biotin and coenzyme R. Science 91:243-244.

47. György P, Rose CS, Eakin RE, Snell EE, Williams RJ. 1941. Egg-white injury as the result of nonabsorption or inactivation of biotin. Science 93:477-478.

48. György P, Kuhn R, Lederert E. 1939. Attempts to isolate the factor (vitamin H) curative of egg white injury. Journal of Biological Chemistry 131:745-759.

49. György P. 1939. The curative factor (vitamin H) for egg white injury, with particular reference to its presence in different foodstuffs and in yeast. Journal of Biological Chemistry 131:733-744.

50. Harris SA, Wolf DE, Mozingo R, Anderson RC, Arth GE, Easton NR, Heyl D, Wilson AN, Folkers K. 1944. Biotin. II. Synthesis of Biotin. Journal of the American Chemical Society 66:1756-1757.

51. Henry Lardy BA, Potter RL, Burris RH. 1949. Metabolic functions of biotin I. The role of biotin in bicarbonate utilization by Lactobacillus arabinosus studied witi C14. Journal of Biological Chemistry 179:721-731.

52. Holmberg A, Blomstergren A, Nord O, Lukacs M, Lundeberg J, Uhlén M. 2005. The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. Electrophoresis 26:501-510.

53. Huang W, Lindqvist Y, Schneider G, Gibson KJ, Flint D, Lorimer G. 1994. Crystal structure of an ATP-dependent carboxylase, dethiobiotin synthetase, at 1.65 A resolution. Structure 2:407-414.

54. Huang W, Jia J, Gibson KJ, Taylor WS, Rendina aR, Schneider G, Lindqvist Y. 1995. Mechanism of an ATP-dependent carboxylase, dethiobiotin synthetase, based on crystallographic studies of complexes with substrates and a reaction intermediate. Biochemistry 34:10985-95.

55. Ifuku O, Miyaoka H, Koga N, Kishimoto J, Haze S-i, Wachi Y, Kajiwara M. 1994. Origin of the carbon atoms of biotin. 13C-NMR studies on biotin biosynthesis in Escherichia coli. European Journal of Biochemistry 220:585-591.

56. Iram SH, Cronan JE. 2006. The ß-oxidation systems of *Escherichia coli* and *Salmonella enterica* are not functionally equivalent. J Bacteriol 188:599-608.

57. Izumi Y, Sato K, Tani Y, Ogata K. 1973. Distribution of 7-Keto-8-aminopelargonic Acid Synthetase in Bacteria and the Control Mechanism of the Enzyme Activity. Agricultural and Biological Chemistry 37:1335-1340.

58. Izumi Y, Morita H, Sato K, Tani Y, Ogata K. 1972. Synthesis of biotin-vitamers from pimelic acid and coenzyme A by cell-free extracts of various bacteria. Biochimica et Biophysica Acta 264:210-213.

59. Izumi Y, Kano Y, Inagaki K, Kawase N, Tani Y, Yamada H. 1981. Characterization of biotin biosynthetic enzymes of Bacillus sphaericus: A dethiobiotin producing bacterium. Agricultural and Biological Chemistry 45:1983-1989.

60. Izumi Y, Morita H, Tani Y, Ogata K. 1973. Partial purification and some properties of 7-Keto-8-aminopelargonic acid synthetase, an enzyme involved in biotin biosynthesis. Agricultural and Biological Chemistry 37:1327-1333.

61. Izumi Y, Morita H, Tani Y, Ogata K. 1974. The Pimelyl-CoA Synthetase Responsible for the First Step in Biotin Biosynthesis by Microorganisms. Agricultural and Biological Chemistry 38:2257-2262.

62. Janota-Bassalik L, Wright LD. 1964. Pimelic acid as a by-product of Azelaic acid degradation by Pseudomomas sp. Nature 204:501-502.

63. Jayasuriya H, Herath KB, Zhang C, Zink DL, Basilio A, Genilloud O, Diez MT, Vicente F, Gonzalez I, Salazar O, Pelaez F, Cummings R, Ha S, Wang J, Singh SB. 2007. Isolation and Structure of Platencin: A FabH and FabF Dual Inhibitor with Potent Broad-Spectrum Antibiotic Activity. Angewandte Chemie International Edition 46:4684-4688.

64. Jencks W. 1963. Preparation and properties of acyl adenylates. Meth Enzymol 6:762-766.

65. Käck H, Sandmark J, Gibson K, Schneider G, Lindqvist Y. 1999. Crystal structure of diaminopelargonic acid synthase: evolutionary relationships between pyridoxal-5′-phosphate-dependent enzymes. Journal of Molecular Biology 291:857-876.

66. Kaneda T. 1967. Fatty Acids in the Genus Bacillus. Iso- and Anteiso-fatty acids as characteristic constituents of lipids in 10 species. Journal of Bacteriology 93:894-903.

67. Knowles JR. 1989. The Mechanism of Biotin-Dependent Enzymes. Annual Review of Biochemistry 58:195-221.

68. Kögl F, Tönnis B. 1936. Über das Bios-Problem. Darstellung von krystallisiertem Biotin aus Eigelb. . Hoppe-Seyler´s Zeitschrift für physiologische Chemie 43:242-273.

69. Krell K, Eisenberg MA. 1970. The Purification and Properties of Dethiobiotin Synthetase. Journal of Biological Chemistry 245:6558-6566.

70. Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessières P, Bolotin A, Borchert S, Borriss R, Boursier L, Brans A, Braun M, Brignell SC, Bron S, Brouillet S, Bruschi CV, Caldwell B, Capuano V, Carter NM, Choi S-K, Codani J-J, Connerton IF, Cummings NJ, Daniel RA, Denizot F, Devine KM, Düsterhöft A, Ehrlich SD, Emmerson PT, Entian KD, Errington J, Fabret C, Ferrari E, Foulger D, Fritz C, Fujita M, Fujita Y, Fuma S, Galizzi A, Galleron N, Ghim S-Y, Glaser P, Goffeau A, Golightly EJ, Grandi G, Guiseppi G, Guy BJ, Haga K, et al. 1997. The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. nature 390:249-256.

71. Lardy HA, Peanasky R. 1953. Metabolic Functions of Biotin. Physiological Reviews 33:560-565.

72. Laurermairs E, Schwarzs E, Oesterheltz D, Reinkeg H, Beyreutherg K, Dimroth P. 1989. The Sodium Ion Translocating Oxaloacetate Decarboxylase of Klebsiella pneumoniae. Sequence of integral memebrane-bound subunits β and γ. Journal of Biological Chemistry 264:14710-14715.

73. Lewis DFV, Wiseman A. 2005. A selective review of bacterial forms of cytochrome P450 enzymes. Enzyme and Microbial Technology 36:377-384.

74. Lezius A, Ringlemann E, Lynen F. Zur biochemischen Funktion des Biotins. IV. die Biosynthese des Biotins. Biochemische Zeitschrift:510-525.

75. Lin S, Hanson RE, Cronan JE. 2010. Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nature Chemical Biology 6:682-688.

76. Lin S, Cronan JE. 2012. The BioC O-Methyltransferase Catalyzes Methyl Esterification of Malonyl-Acyl Carrier Protein, an Essential Step in Biotin Synthesis. Journal of Biological Chemistry 287:37010-37020.

77. Lin S, Cronan JE. 2011. Closing in on complete pathways of biotin biosynthesis. Molecular BioSystems 7:1811-1811.

78. Loftfield RB, Eigner EA. 1966. The specificity of enzymic reactions. Aminoacyl-soluble RNA ligases. Biochim Biophys Acta 130:426-48.

79. Loftfield RB, Vanderjagt D. 1972. The frequency of errors in protein biosynthesis. Biochem J 128:1353-6.

80. Lynen F. 1967. The Role of Biotin-Dependent Carboxylations in Biosynthetic Reactions. Biochemical Journal 102:381-400.

81. Makris TM, Davydov R, Denisov IG, Hoffman BM, Sligar SG. 2002. Mechanistic enzymology of oxygen activation by the cytochromes P450. Drug Metab Rev 34:691-708.

82. Manandhar M, Cronan JE. 2013. Proofreading of Noncognate Acyl Adenylates by an Acyl-Coenzyme A Ligase. Chemistry & Biology 20:1441-1446.

83. Marquet A, Tse Sum Dui B, Florentin D. 2001. Biosynthesis of Biotin and Lipoic Acid, p 51-101, vol 61.

84. Martin N, Christensen QH, Mansilla MC, Cronan JE, de Mendoza D. 2011. A novel two-gene requirement for the octanoyltransfer reaction of Bacillus subtilis lipoic acid biosynthesis. Molecular Microbiology 80:335-349.

85. Martinis SA, Boniecki MT. 2010. The balance between pre- and post-transfer editing in tRNA synthetases. FEBS Lett 584:455-9.

86. Meir A, Bayer EA, Livnah O. 2012. Structural Adaptation of a Thermostable Biotin-binding Protein in a Psychrophilic Environment. Journal of Biological Chemistry 287:17951-17962.

87. Melville DB, Moyer AW, Hofmann K, Du Vigneaud V. 1942. The structure of biotin: the formation of thiophenevaleric acid from biotin. Journal of Biological Chemistry 146:487-492.

88. Mueller JH. 1937. Pimelic acid as a growth accessory factor for a strain of the Diphtheria bacillus. Science 85:502-503.

89. Mueller JH. 1937. Pimelic acid as a growth accessory for the Diphtheria Bacillus. Journal of Biological Chemistry 119:121-131.

90. Mueller JH. 1937. Studies on cultural requirements of bacteria. Pimelic acid as a growth stimulant for C. Diphtheriae. Journal of Bacteriology 34:163-178.

91. Mueller JH. 1937. Pimelic acid as a growth accesory factor for the strain of the Diptheria Bacillus. Science 85:1935-1936.

92. Munro AW, Lindsay JG. 1996. Bacterial cytochromes P-450. Molecular Microbiology 20:1115-1125.

93. Nakano MM, Zuber P. 1998. Anaerobic growth of a "strict aerobe" (Bacillus subtilis). Annual Review of Microbiology 52:165-190.

94. O'Keefe SJ, Knowles JR. 1986. Biotin-Dependent Carboxylation Catalyzed by Transcarboxylase Is a Stepwise Process. Biochemistry 25:6077-6084.

95. Ogata K, Tochikura T, Osugi M, Iwahara S. 1966. Fatty Acid Metabolism in Microorganisms Part I. Production of pimelic acid from azelaic acid. Agricultural and Biological Chemistry 30:176-180.

96. Ogata K. 1970. Microbial synthesis of dethiobiotin and biotin. Methods in Enzymology 18:390-394.

97. Ohsugi M, Miyauchi K, Inoue Y. 1984. Pimelic Acid as a Degradation Product of Azelaic Acid by Yeasts. Agricultural and Biological Chemistry 48:1881-1882.

98. Ohsugi M, Miyauchi K, Inoue Y. 1983. Pimelic Acid Determination with Resting Cell System of Bacillus sphaericus AKU 0218. Agricultural and Biological Chemistry 47:1649-1650.

99. Ohsugi M, Miyauchi K, Tachibana K, Nakao S. 1988. Formation of a biotin precursor, pimelic acid, in yeasts from C18 fatty acids. Journal of Nutritional Science and Vitaminology 34:343-352.

100. Omura T, Sato R. 1964. The Carbon Monoxide-binding Pigment of Liver Microsomes II. Solubilization, purification, and properties. Journal of Biological Chemistry 239:2379-2385.

101. Pai CH. 1971. Biosynthesis of Biotin: Synthesis of 7, 8- Diaminopelargonic Acid in Cell-Free Extracts of Escherichia coli. Journal of Bacteriology 105:793-800.

102. Pai CH, McLaughlin GE. 1969. Uptake of pimelic acid by Escherichia coli and Pseudomonas denitrificans. Canadian Journal of Microbiology 15:809-810.

103. Pauling L. 1958. Festschrift für Prof Dr. Arthur Stoll. Birkhauser Verlag, Basel, Switzerland.

104. Perkins JB, Bower S, Howitt CL, Yocum RR, Pero J. 1996. Identification and characterization of transcripts from the biotin biosynthetic operon of Bacillus subtilis. Journal of Bacteriology 178:6361-6365.

105. Ploux O, Marquet A. 1996. Mechanistic Studies on the 8-Amino-7-Oxopelargonate Synthase, a Pyridoxal-5'-Phosphate-Dependent Enzyme Involved in Biotin Biosynthesis. European Journal of Biochemistry 236:301-308.

106. Ploux O, Marquet A. 1992. The 8-amino-7-oxopelargonate synthase from Bacillus sphaericus . Purification and preliminary characterization of the cloned enzyme overproduced in Escherichia coli. Biochemical Journal 283:327-331.

107. Ploux O, Soularue P, Marquet A, Gloeckler R, Lemoine Y. 1992. Investigation of the first step of biotin biosynthesis in Bacillus sphaericus . Purification and characterization of the pimeloyl-CoA synthase, and uptake of pimelate. Biochemical Journal 287:685-690.

108. Porebski PJ, Klimecka M, Chruszcz M, Nicholls RA, Murzyn K, Cuff ME, Xu X, Cymborowski M, Murshudov GN, Savchenko A, Edwards A, Minor W. 2012. Structural characterization of Helicobacter pylori dethiobiotin synthetase reveals differences between family members. FEBS Journal 279:1093-1105.

109. Priest FG. 1993. Systematics and Ecology of Bacillus. In Bacillus Subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics. p.3-16. American Society of Microbiology.

110. Rodionov Da. 2002. Conservation of the Biotin Regulon and the BirA Regulatory Signal in Eubacteria and Archaea. Genome Research 12:1507-1516.

111. Rolfe B, Eisenberg MA. 1968. Genetic and biochemical analysis of the biotin loci of Escherichia coli K-12. Journal of bacteriology 96:515-24.

112. Rossmann M, Argos P. 1978. The taxonomy of binding sites in proteins. Molecular and Cellular Biochemistry 21:161-182.

113. Sanyal I, Lee S-L, Flint DH. 1994. Biosynthesis of pimeloyl-CoA, a biotin precursor in Escherichia coli, follows a modified fatty acid synthesis pathway: 13C-labeling studies. Journal of the American Chemical Society 116:2637-2638.

114. Schall OF, Suzuki I, Murray CL, Gordon JI, Gokel GW. 1998. Characterization of acyl adenyl anhydrides: Differences in the hydrolytic rates of fatty acyl-AMP and aminoacyl-AMP derivatives. Journal of Organic Chemistry 63:8661-8667.

115. Schujman GE, Choi KH, Altabe S, Rock CO, De Mendoza D. 2001. Response of Bacillus subtilis to cerulenin and acquisition of resistance. Journal of Bacteriology 183:3032-3040.

116. Sekiguchi J, Takada N, Okada H. 1975. Genes affecting the productivity of alpha-amylase in Bacillus subtilis Marburg. Journal of Bacteriology 121:688-694.

117. Shapiro MM, Chakravartty V, Cronan JE. 2012. Remarkable Diversity in the Enzymes Catalyzing the Last Step in Synthesis of the Pimelate Moiety of Biotin. PLoS ONE 7:e49440-e49440.

118. Spizizen J. 1958. Transformation of biochemically deficient strains of Bacillus subtilis by deoxyribonucleate. Proceedings of the National Academy of Sciences 44:1072-1078.

119. Stok JE, De Voss JJ. 2000. Expression, Purification, and Characterization of BioI: A Carbon–Carbon Bond Cleaving Cytochrome P450 Involved in Biotin Biosynthesis in Bacillus subtilis. Archives of Biochemistry and Biophysics 384:351-360.

120. Swick RW, Wood HG. 1960. The role of transcarboxylation in propionic acid fermentation. Biochemistry 46:28-41.

121. Tong L. 2013. Structure and function of biotin-dependent carboxylases. Cellular and Molecular Life Sciences 70:863-891.

122. Ugulava NB, Gibney BR, Jarrett JT, Jarrett JT. 2001. Biotin Synthase Contains Two Distinct Iron−Sulfur Cluster Binding Sites: Chemical and Spectroelectrochemical Analysis of Iron−Sulfur Cluster Interconversions †. Biochemistry 40:8343-8351.

123. Vagner V, Dervyn E, Ehrlich SD. 1998. A vector for systematic gene inactivation in Bacillus subtilis. Microbiology 144:3097-3104.

124. Van Arsdell SW, Perkins JB, Yocum RR, Luan L, Howitt CL, Prasad Chatterjee N, Pero JG. 2005. Removing a bottleneck in theBacillus subtilis biotin pathway: BioA utilizes lysine rather thanS-adenosylmethionine as the amino donor in the KAPA-to-DAPA reaction. Biotechnology and Bioengineering 91:75-83.

125. Vanaman TC, Wakil SJ, Hill RL. 1968. The complete amino acid sequence of the acyl carrier protein of Escherichia coli. Journal of Biological Chemistry 243:6420-6431.

126. Wang J, Kodali S, Lee SH, Galgoci A, Painter R, Dorso K, Racine F, Motyl M, Hernandez L, Tinney E, Colletti SL, Herath K, Cummings R, Salazar O, Gonzalez I, Basilio A, Vicente F, Genilloud O, Pelaez F, Jayasuriya H, Young K, Cully DF, Singh SB. 2007. Discovery of platencin, a dual FabF and FabH inhibitor with in vivo antibiotic properties. Proceedings of the National Academy of Sciences 104:7612-7616.

127. Wang M, Moynié L, Harrison PJ, Kelly V, Piper A, Naismith JH, Campopiano DJ. 2017. Using the pimeloyl-CoA synthetase adenylation fold to synthesize fatty acid thioesters. Nature Chemical Biology doi:10.1038/nchembio.2361.

128. Webster SP, Alexeev D, Campopiano DJ, Watt RM, Alexeeva M, Sawyer L, Baxter RL. 2000. Mechanism of 8-Amino-7-oxononanoate Synthase: Spectroscopic, Kinetic, and Crystallographic Studies Biochemistry 39:516-528.

129. Wildiers E. 1901. Nouvelle substance indispensable au developpement de la levure. La cellule 18:313-316.

130. Yadavalli SS, Ibba M. 2012. Quality control in aminoacyl-tRNA synthesis its role in translational fidelity. Adv Protein Chem Struct Biol 86:1-43.

131. Yao J, Rock CO. 2016. Bacterial fatty acid metabolism in modern antibiotic discovery. Biochim Biophys Acta doi:10.1016/j.bbalip.2016.09.014.

132. Zeigler DR, Pragai Z, Rodriguez S, Chevreux B, Muffler A, Albert T, Bai R, Wyss M, Perkins JB. 2008. The Origins of 168, W23, and other Bacillus subtilis legacy strains. Journal of Bacteriology 190:6983-6995.